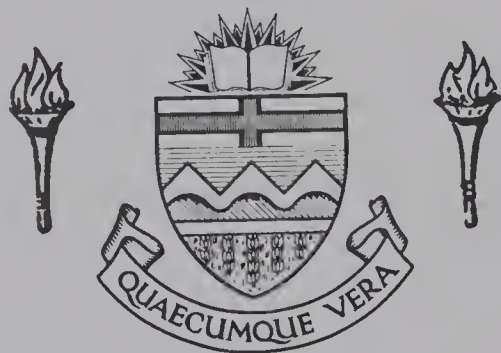


For Reference

NOT TO BE TAKEN FROM THIS ROOM

Ex LIBRIS
UNIVERSITATIS
ALBERTAENSIS



THE UNIVERSITY OF ALBERTA
A GRADE SIX CASE STUDY OF AN
ILLUSTRATIVE TIME LINE'S
EFFECTIVENESS

by



FLOYD M. WOLLUM

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF EDUCATION

DEPARTMENT OF ELEMENTARY EDUCATION

EDMONTON, ALBERTA
(FALL), 1970

Thesis
1878
287

UNIVERSITY OF ALBERTA
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend
to the Faculty of Graduate Studies for acceptance, a thesis
entitled A GRADE SIX CASE STUDY OF AN ILLUSTRATIVE TIME LINE'S
EFFECTIVENESS submitted by Floyd M. Woilum in partial fulfilment
of the requirements for the degree of Master of Education.

ACKNOWLEDGEMENTS

The writer expresses his thanks to his advisor, Dr. Joseph Kirman, who, through his well thought-out comments and criticisms, has helped to improve the conciseness and clarity of this thesis.

Thanks also to Dr. George Cathcart for his patience and competence in supplying valuable assistance with the statistical analyses carried out during this study.

Appreciation is expressed for the interest and advice offered by Dr. E.C. Wilson and Dr. D.F. Kellum, members of the committee.

In addition, acknowledgement is given for the courtesy and help offered by Mr. Robert Carter, assistant to the Superintendent of the Grande Prairie Public School District, and the grade six teachers on his staff.

Finally, this writer wishes to thank his wife, Bridget, who acted as typist, for her great persistence and dedication. Also, appreciation is extended to any other people who have helped to bring this thesis to completion.

ABSTRACT

Within this study an attempt has been made to help answer the following question: Is the effect of an illustrative time line with a historical narrative superior to the historical narrative alone in terms of pupil understanding of the chronological relationships in the narrative?

The study was carried out during May, 1970, in the Grande Prairie Public School District, Alberta, at the grade six level, using 214 subjects assigned randomly to two treatment groups. Treatment materials for the experimental and control groups, respectively, consisted of a 750 word historical narrative with an illustrative time line and the historical narrative by itself. Eight dated events were incorporated into the narrative as well as shown on the time line. Pupils were given one hour to read the treatment materials and do the test.

Analysis of the data in terms of I.Q. groups, reading comprehension groups, sex, and chronological and mental ages yielded the following findings:

1. The effect of the illustrative time line with the historical narrative was superior to the effect of the

narrative alone:

- a. generally
- b. at high and high average I.Q. levels but not at average and low I.Q. levels.
- c. at low, average, and high reading levels.
- d. for boys or girls.

2. There was no interaction between chronological or mental age and treatment. Also, increased chronological or mental ages were not good predictors of increased understanding of chronological relationships.

TABLE OF CONTENTS

Chapter		Page
1.	THE PROBLEM, ITS NATURE AND SIGNIFICANCE	1
	INTRODUCTION	1
	THE PROBLEM	5
	HYPOTHESES	7
	DEFINITIONS	8
	SIGNIFICANCE	9
2.	REVIEW OF THE RELATED RESEARCH LITERATURE	12
	RESEARCH SPECIFICALLY ON TIME LINES	12
	RELATED STUDIES ON ILLUSTRATIVE MATERIALS	16
	RESEARCH ON TIME AND CHRONOLOGY	16
	RELATING THE HYPOTHESES TO THE LITERATURE	31
3.	METHODOLOGY	45
	DESIGN OF THE EXPERIMENT	45
	THE DEPENDENT AND INDEPENDENT VARIABLES	50
	SAMPLE	51
	VALIDITY AND RELIABILITY OF THE TEST	
	INSTRUMENTS USED	52
	Instruments Used	52
	Reading comprehension test	53
	Description	53
	Assessment of validity and reliability	55

Use in the present study	56
California Short-Form Test of Mental Maturity	57
Description	58
Assessment of validity and reliability	60
Use in the present study	62
Chronological Relationships Test	63
Description	63
Validity	64
Reliability	68
Use in the present study	69
DATA GATHERING PROCEDURES	69
Gathering Data on the Independent Variables	69
Applying Control and Experimental Treatments	70
Preliminaries	70
Directions for utilizing answer sheets	71
Random distribution of treatment materials	72
Reading the treatment materials	72
Instructions for doing the Chrono- logical Relationships Test	72
Doing the Chronological Relationships Test	73
Scoring the Chronological Relationships Test	73
NULL HYPOTHESES TESTED	74
STATISTICAL PROCEDURES	75

Chapter	Page
The T-test	76
Assumption of normality	76
Homogeneity of variance	76
Continuity and equal intervals as measures	76
Test for Interaction	79
SUMMARY	80
4. RESULTS	82
NULL HYPOTHESES RELATING TO INTELLIGENCE QUOTIENT	82
NULL HYPOTHESES RELATING TO READING COMPREHENSION LEVELS	84
NULL HYPOTHESIS RELATING TO AGE	86
NULL HYPOTHESES RELATING TO SEX	87
SUMMARY	88
5. CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS	90
REVIEW OF THE PROBLEM AND PROCEDURES	90
CONCLUSIONS	91
DISCUSSION	93
LIMITATIONS	98
IMPLICATIONS	99
RECOMMENDATIONS FOR RESEARCH	100
BIBLIOGRAPHY	102
APPENDICES	106

Chapter 1

THE PROBLEM, ITS NATURE AND SIGNIFICANCE

INTRODUCTION

Although a sense of time and chronology is helpful in all of the social studies, it is indispensable to the study of history, which is the record of change through time.¹ What is now happening may be considered as partly the result of what has happened, and the study of human activity in time can make the abstract idea that events have consequences specific and understandable. However, the research indicates that it would be naive to expect pupils in elementary schools, even with systematic instruction, to display characteristics of one who has mature sense of time relationships.²

¹K.C. Friedman and V.A. Marti, "A Time Comprehension Test," Journal of Educational Research, XXXIX (Sept., 1945), p. 68; A.W. Spieseke, "Developing a Sense of Time and Chronology," Skill Development in the Social Studies, Thirty-third Yearbook of the National Council for the Social Studies (Washington: N.C.S.S., 1963), p. 174.

²A. Flickinger and K.J. Rehage, "Building Time and Space Concepts," Improving the Teaching of World History, Twentieth Yearbook of the National Council for the Social Studies (Washington: N.C.S.S., 1949), p. 110.

A person who has acquired a mature sense of time relationships would be able to use correctly the terms involved in our system of reckoning time. He would be able to interpret correctly, within reasonable limits, time words of varying degrees of definiteness. He would habitually inquire "When?" in his study of historical events, movements, and personalities. He would answer this question by fitting newlyacquired information into a basic framework provided by a few key dates. He would be able to use time lines to facilitate the development of his sense of chronology. He would be able to arrange a series of related events in chronological order. He would be able to bridge the gap between the present and the period he is studying in such a way as to facilitate his identification with the people of that period.

But a mature sense of time involves more than the ability to see important events merely as discrete "points" on a time line. It would require also that one see such events as phenomena deeply rooted in their own past and extending their influence well beyond the time of their occurrence. This "historical" view of time depends upon the basic ideas of duration and continuity.³

Flickinger and Rehage go on to give suggestions for teaching a sense of time relationships. Among these they suggest that pupils should have pratice in drawing generalizations not only from their own time lines but from time lines prepared by others as well.⁴

Such time lines are sometimes used to present or to illustrate information about time and chronology presented in historical passages in social studies textbooks. In addition, it has been claimed by some that the time line is a useful tool for helping children to deal with time and chronology.⁵

³ Ibid., pp. 108-09.

⁴ Ibid., p. 111.

⁵ K.C. Friedman, "Time Concepts of Elementary-School Children," (Hereinafter referred to as Friedman I),

It has been pointed out in two studies by Davis⁶ that few studies have actually investigated the specific role of the time line in the learning of chronological relationships. In Davis' earlier study at the grade eleven and twelve level results indicated that a correct time line was most useful for high school students of a relatively high I.Q. level in learning chronological relationships in a textual narrative. Such a time line was found beneficial to such students only in comparison with a situation in which a time line was not present with the narrative.⁷

In a second study by Davis, at the junior high school level, he found no significant difference between two groups in scores on a chronological relationships test. One group studied only a narrative while the other had the narrative as well as a relevant complete time line. In both treatments

Elementary School Journal, XXXIV (February, 1944), 342; Flickinger and Rehage, op. cit., p. 111; E.B. Wesley and M.A. Adams, Teaching Social Studies in Elementary Schools, (rev. ed.; Boston: D.C. Heath and Co., 1952), p. 304; W.L. Chase, A Guide for the Elementary Social Studies Teacher (Boston: Allyn and Bacon, 1968), pp. 80-81; Spieseke, op. cit., p. 185; J. Jarolimek, Social Studies in Elementary Education (3rd. ed.; Toronto: Collier-Macmillan, 1967), p. 345; R.C. Preston, Teaching Social Studies in Elementary School (3rd. ed.; Toronto: Holt, Rinehart and Winston, 1968), p. 316.

⁶O.L. Davis, L.C. Hicks and N.D. Bowers, "The Usefulness of Time Lines in Learning Chronological Relationships in Text Materials," Journal of Experimental Education, XXXIV (Spring, 1966), 22-25; O.L. Davis, "Usefulness of a Time Line - With Historical Text," Educational Leadership, XXV (May, 1968), p. 752.

⁷Davis, Hicks, and Bowers, op. cit., pp. 22-25.

pupils at higher grade or I.Q. levels demonstrated higher achievement on the test. Unlike the earlier study, high I.Q. students did not achieve better with the aid of the time line than without it.⁸

Both of the preceeding studies were above the elementary school level. No studies specifically on the effectiveness of time lines at the elementary school level have been done up to now so far as this investigator can determine from the articles listed in all volumes of the Education Index⁹ and all available volumes of ERIC¹⁰ as well as from leads provided in the bibliographies of the studies which were investigated as part of the review of the literature in the present study.¹¹ However, there have been studies of the effect of instruction about time and chronology incorporating time lines as one of several aids.¹²

⁸O.L. Davis, op. cit., p. 752.

⁹Education Index, (New York: H.W. Wilson Co.), Jan. 1929-May, 1970 inclusive.

¹⁰Educational Resources Information Center, (Washington: U.S. Office of Education), Jan., 1956-June, 1970 inclusive.

¹¹See Chapter 2.

¹²K.C. Friedman, "Time Concepts of Junior and Senior High School Pupils and of Adults," (Hereinafter referred to as Friedman II), School Review, LII (April, 1944), pp. 233-38; F. Pistor, "How Time Concepts Are Acquired By Children," (Hereinafter referred to as Pistor I), Educational Method, XX (November, 1940), pp. 107-12; Friedman and Marti, op. cit., pp. 62-68.

THE PROBLEM

There seems to be a measure of agreement on the assumption that before the fifth or sixth grade most children have little or no sense of chronology.¹³ However, we do not know whether time lines used in grade five or six social studies textbooks add to the pupils' understanding of chronological relationships described in written form. To help fill this gap in knowledge about time lines a study at grade six level has been carried out by this investigator.

Following the lead provided by Davis, the present study attempted to answer this question: Is the effect of an illustrative time line with a historical narrative superior to the historical narrative alone in terms of pupil understanding of the chronological relationships in the narrative?

The degree of pupil understanding of chronological relationships has been measured by the degree of achievement on a chronological relationships test consisting of questions

¹³D.H. Russel, Children's Thinking (Toronto: Ginn and Co., 1956), pp. 133-34; W.L. Chase, "American History in the Middle Grades," (Hereinafter referred to as Chase, "American History"), Interpreting and Teaching American History, Thirty-first Yearbook of the National Council for the Social Studies (Washington: N.C.S.S., 1961), p. 333; Spieseke, op. cit., p. 177; Jarolimek, op. cit., p. 345; Preston, op. cit., p. 150.

about the order of events and durations of time described in the narrative.¹⁴

The following items, condensed from the literature and elaborated on in the next chapter, might affect hypotheses about the effect of an illustrative time line at grade six level.

First, the value of time lines as illustrations to aid learning from texts is doubtful unless intelligence is considered.¹⁵ It appears that at higher intelligence levels understanding of time concepts and ability to use a time line improve.¹⁶

Second, reading level has apparently not been investigated as a factor affecting utilization of time lines.¹⁷

Third, apparently maturation aids in the development

¹⁴See Chapter 3, pp. 82-88 and Appendices A and B.

¹⁵Davis, et al, op. cit., pp. 22-25; Davis, op. cit., pp. 750-53; J.G. Burdick, "A Study of Cross-Section Drawings as Technical Illustrations in High School Textbooks," (unpublished Ph.D. dissertation, Syracuse Univ., 1959), pp. 43-44; W.A. Miller "Reading With and Without Pictures," Elementary School Journal, XXXVIII (May, 1938), pp. 678-82; M.D. Vernon, "The Value of Pictorial Illustration," British Journal of Educational Psychology, XXIII (1953), pp. 182-83.

¹⁶W.O. Newman, "Childrens' Understanding of Time Duration" (unpublished Master's thesis, University of Alberta, Edmonton, Canada, 1967), p. 96; Friedman II, op. cit., pp. 237-38; Davis, op. cit., p. 752.

¹⁷See review of the literature in Chapter 2, pp. 33-34.

of a sense of time and chronology.¹⁸

Fourth, it seems likely that sex may help to determine children's knowledge of and ability to understand time and chronology.¹⁹

HYPOTHESES

In terms of pupil achievement on a chronological relationships test the effect of a historical narrative with a time line;

1. is superior to the effect of the narrative alone

¹⁸E.C. Oakden and M. Sturt, "The Development of the Knowledge of Time in Children," British Journal of Psychology, XII (April, 1922), pp. 334-36; M.G. Kelty, "Time Expressions Comprehended by Children in the Elementary School," Elementary School Journal, XXV (April, 1925), pp. 615-18; M.L. Harrison, "The Nature and Development of Concepts of Time Among Young Children," Elementary School Journal, XXXIV (March, 1934), pp. 502-14; N.C. Bradley, "The Growth of the Knowledge of Time in Children of School Age," British Journal of Psychology, XXXVIII (1947), pp. 77-78; M.G. Callahan, "The Construction and Evaluation of a Test in Time Relationships for Grades V, VI, VII, and VIII" (unpublished Ed. M. Thesis, Boston University School of Education, 1952) cited by W.L. Chase, "American History," op. cit., pp. 337-38; M. Farrell, "Understanding of Time Relations of Five-, Six-, and Seven-year-old Children of High I.Q.," Journal of Educational Research, XLVI (April, 1953), pp. 592-94; C.G. Gill, "Interpretations of Indefinite Expressions of Time," Social Education, XXVI (December, 1962), pp. 455-56; Newman, op. cit., pp. 96-97.

¹⁹Friedman II, op. cit., pp. 237-38; Farrell, op. cit., pp. 592-93; Newman, op. cit., p. 96; Callahan, op. cit., pp. 337-38.

- at the high intelligence level.²⁰
2. is the same as the effect of the narrative alone at low, average, or high reading comprehension levels.²¹
 3. increases equally with the effect of the narrative alone as chronological age increases.
 4. is the same as the effect of the narrative alone for boys or girls.

DEFINITIONS

Terms used in the hypotheses are defined here as they are used in the investigation.²²

Maturation. Refers to changes in an individual's intellectual or mental capacity due to increased chronological age.

Chronological relationships test. A test incorporating questions about the order and duration of events described in the historical narrative.

Illustrative time line. A line upon which events are arrayed on a continuum which indicates both the order of and differences in elapsed time between all of the dated events

²⁰As measured by the California Short Form Test of Mental Maturity, Level 2.

²¹As measured by the grade six reading comprehension section of the Canadian Tests of Basic Skills, Form 1.

²²See Chapter 3 for discussion of validity, reli-

from an accompanying written historical narrative.

Historical narrative. A story or article containing events dated according to our Gregorian Calendar.

Reading level. Refers to low, average, or high reading achievement level as determined by score on the grade six reading comprehension test from the Canadian Tests of Basic Skills, Form 1.

Intelligence. Refers to high, high average, average, or low intelligence quotient as determined by the California Short-Form Test of Mental Maturity, Level 2.

Age. Present chronological age of the pupil in months.

Chronology. The science that deals with measuring time by regular divisions and that assigns to events their proper dates.²³

SIGNIFICANCE

In school text materials historical narratives are often accompanied by time lines. This is especially the case when chronological relationships are deemed critical.

The usefulness of such time lines in promoting pupil learning has been advanced without a substantial research

ability, cut-off points for grouping, and procedures for administration of tests mentioned in definitions.

²³Webster's Seventh New Collegiate Dictionary (1969) pp. 148-49.

base. In two previous studies²⁴ the usefulness of a time line with a historical text has been examined empirically. The first study found that a time line did help high intelligence high school students to learn more while the second study found that a time line apparently did not help junior high school students to learn more.²⁵

By investigating the usefulness of time lines in upper elementary school this study tested time line usefulness with less mature students, thus extending the two previous studies without their procedural omissions.²⁶

If this study shows that the illustrative time line is not useful then the value of the inclusion of such time lines in elementary school textbooks should merit further study by other investigators. Conditions under which time lines merit inclusion or whether they should be included at all could be investigated. On the other hand, if the time line proves to be useful then this study will lend support to the inclusion of time lines as illustrations in textual materials.

Hopefully, the findings and recommendations of this study will lead to investigation of ways in which time lines

²⁴Davis, et al, op. cit., pp. 23-25; Davis, op. cit., pp. 751-52.

²⁵See criticism of their procedures in Chapter 2, pp. 14-15.

²⁶See Chapter 2, pp. 14-15 and p. 32.

can be most effectively used if they are to be used in elementary schools.

Chapter 2

REVIEW OF THE RELATED RESEARCH LITERATURE

In the following review of the literature three kinds of research have been examined:

First, the research specifically on time lines has been reported.

Second, some studies which relate to the usefulness of illustrations with written text have been reviewed.

Third, research on various aspects of children's understanding of time and chronology has been reported.

Information derived from the literature is used to justify the hypotheses which were set down in chapter one.¹ Following the review of the research literature hypotheses one to four will each be restated and related to relevant research.

RESEARCH SPECIFICALLY ON TIME LINES

From the literature available it appears that two

¹ See pp. 7-8.

studies specifically on the usefulness of time lines have been done. The first of these time line studies² assessed the usefulness of a time line in learning chronological relationships in text materials. A sample of 81 pupils from a large world history class at grade eleven and twelve level was selected randomly. The purpose was primarily to assess the value of a time line as a device to illustrate a sequence of seven events mentioned in a 700 word historical narrative.

Three conditions consisting of the narrative plus a correct time line, the narrative plus and incorrect time line, and the narrative by itself were presented to three equal-sized, randomly chosen, groups. Following study of the materials they were removed and a test on the chronological relationships incorporated in the narrative plus correct time line was administered.

From the study two major conclusions were forthcoming. First, high intelligence students benefitted from the correct time line when compared with high intelligence students with no time line. Second, the correct time line did not appear to increase student learning when contrasted with the incorrect time line. As a result, the study raised questions about the usefulness of time lines as illustrations of printed verbal narratives unless intelli-

²Davis, et al, op. cit., pp. 22-25.

gence is considered.

A subsequent study by Davis³ involved 538 junior high school pupils. Two treatments consisting of a narrative with no time line and a narrative with a complete illustrative time line were administered. Pupils were assigned randomly to one or the other treatment.

There was no significant difference between treatments on a test on the chronological relationships embodied in the narrative and time line. Furthermore, there was no significant difference between scores on the treatments when controlling for I.Q. and grade levels. Thus, like his first study, Davis' findings failed to support the uncritical assumption of time lines' usefulness with historical text, at least for junior and senior high school pupils.

At this point a major criticism of the main time line studies is in order. This criticism reveals some possible unforeseen consequences of the procedures used.

In both studies the students were given an opportunity to study the materials which had been assigned. However, the materials were taken away while the test was administered making the effect discussed in the following paragraph likely.

On a test about a sequence of either seven or eight

³Davis, op. cit., p. 752.

events embedded in either 700 or 1150 word narrative studied for either 10 or 15 minutes respectively⁴ there would likely be a great lowering of scores due to memory failure. Such a loss could have reduced scores to dependence on the number of discrete bits of information retainable in the memory, regardless of treatment. Thus, the time line might aid learning of the chronological relationships in the narrative while the test might assess only the students' immediate memory span regardless of whether their understanding of the narrative was better with or without the time line.

To overcome this likely procedural defect the present study is not on learning in the sense of memorizing material such as dates found in the narrative and on the line. Instead, as stated previously,⁵ the emphasis is on trying to determine if an illustrative time line helps students to understand the chronological relationships incorporated in the narrative and time line. Second, since the subjects were not expected to learn these chronological relationships as in Davis' study, the written materials were not taken away while the students answered the test questions. In such a way, dependence on memory has been minimized in the present study.

⁴Davis et al., op. cit., p. 23; Davis, op. cit., pp. 751-52.

⁵See p. 5.

RELATED STUDIES ON ILLUSTRATIVE MATERIALS

It is useful to point out that the findings of Davis, Hicks, and Bowers, and Davis are consistent with the findings of some other related studies dealing with illustrative materials. For example, Miller, in a study of reading with and without pictures in primary grades,⁶ concluded that it is not possible to infer that illustrations have an effect on reading comprehension. Similarly, Burdick⁷ reported no significant difference in comprehension of science information when a textual narrative was accompanied by drawings and when it was studied without illustration. Neither Miller nor Burdick took intelligence into consideration. However, Vernon⁸, in a study of the comprehension of textbook materials with and without graphs at the high school level, observed that intelligence of the reader is probably the most important condition in understanding graphical material.

RESEARCH ON TIME AND CHRONOLOGY

None of the following studies has dealt primarily

⁶Miller, op. cit., pp. 678-82.

⁷Burdick, op. cit., p. 37.

⁸Vernon, op. cit., pp. 186-87.

with time lines. Generally, they have been concerned with the nature and development of children's concepts of time and chronology. However, information about the effects of variables such as intelligence, reading level, age, sex, and socio-economic status which might be pertinent to the directive hypotheses of the present study has been looked at in the following review. After the literature has been reviewed the items relating to the hypotheses of the present study will be discussed and summarized so as to show what justification there is for those hypotheses.

Most of the following studies were concerned with findings on two or more variables, such as I.Q. or grade, affecting development of concepts of time and chronology. Thus, since the studies often overlap each other they are reviewed here in chronological order.

Early work by Oakden and Sturt gave a historical and philosophical background⁹ as well as experimental evidence¹⁰ about time concepts. This evidence involved the use of eight types of tests; the easier types were used for 100 children aged four to ten years and the harder types for groups of from 28 to 75 children aged eight to thirteen years. With the younger group the easiest questions were "What is your

⁹M. Sturt, The Psychology of Time (London: Routledge and Kegan Paul, 1922).

¹⁰Oakden and Sturt, op. cit., pp. 309-36.

age?" and "Will you come to school on Saturday?" whereas the hardest questions were about duration of time, such as "How long would it take you to walk around this room?"

In an absurdities test supposedly involving Caesar's conquest of Britain most of the children discovered the absurdity of gunfire and of Caesar's grandson still being alive, but many missed the 30th of February and the statement that 58 B.C. was three years after 55 B.C. Other tests involved such items as the temporal order of historical characters and judging about lengths of time.

Results of the test led Oakden and Sturt to believe that at approximately ten years of age the child begins to acquire some of the concepts necessary for the study of history. Some other general conclusions about the growth of children's knowledge of conventional time seem pertinent also. Such growth appears to be a slow process starting before four years of age and arriving nearly at the adult level at about 13 or 14 years of age. The most important period in its development seems to be about the age of 11, when there appears to be a rapid improvement in all types of time knowledge.

In dealing with historical epochs those which are most remote from our own time are most readily distinguished. This seems to be due to the fact that the earliest distinction is between the present and a past which is mainly negatively characterized. In this past subdivisions are at first ignored and they are mainly attended to after

the age of 11 years.¹¹

Another early study of children's time sense was done by Kelty in 1925.¹² She investigated children's understanding of indefinite expressions of time such as "recently" and "in the Middle Ages". Her work showed that there is a steady increase in the number of pupils from grades four to six who find it easier to understand definite rather than indefinite expressions of time.

For example, younger children understand and use indefinite expressions such as "a long time ago," "before I was born" or "this morning." However, as they increase in age they become able to understand and use more definite expressions such as "in 1492," "twenty-five years ago" or "at 7:00 a.m." Understanding of indefinite expressions of time does not increase a great deal with age because of the varying interpretations put on indefinite expressions such as "a long time ago."

A further study, focusing on children's understanding of fifty commonly used time terms such as "before," "minute," and "month" was conducted by Harrison in 1934.¹³ One hundred and sixty children from kindergarten to grade three were

¹¹Ibid., p. 334.

¹²Kelty, op. cit., pp. 607-18.

¹³Harrison, op. cit., pp. 513-14.

subjects and the scoring device was a seven point scale of the degree of comprehension of the chosen words.

Data gathered indicated that there is steady growth in the comprehension of time terms grade by grade. In addition she found the correlation between grasp of concepts of time and mental age to be $+ .70 \pm .03$; between grasp of concepts of time and chronological age $+ .58 \pm .05$. On this basis Harrison postulated that concepts of time develop in closer agreement with inner maturation, as indicated by mental maturity, than with training and experience, as indicated by chronological age and grade placement.

In 1940 a study on the effect of specific training on the acceleration of development of concepts of time was carried out by Pistor.¹⁴ He set up an experiment to see if history instruction assists the development of time concepts. Two groups were used in the study, each made up of three hundred and twenty grade six pupils. The first group contained children who had worked through separate courses in history and geography. In the second group the children had done only geography courses with history used incidentally.

At the beginning of grade six a battery of time tests constructed by Pistor¹⁵ was administered. Scores for both

¹⁴Pistor I, op. cit.; pp. 107-12.

¹⁵F. Pistor, "Measuring the Time Concepts of Children," (Hereinafter referred to as Pistor II), Journal of Educational Research, XXXIII (December, 1939), 293-300.

groups were nearly identical. Following a year of instruction in which the "history" group was taught by methods emphasizing chronology and the "geography" group with only incidental mention of chronology the test was again administered. Mean scores were almost identical for the two groups again. From this Pistor concluded that maturation rather than training, is the main factor in time concept development.

Friedman, in a 1944 study,¹⁶ constructed two separate tests in order to sample elementary-school children's understanding of conventional time terms. Six hundred and ninety-seven pupils were tested. The first test, administered from kindergarten to grade three, was designed to sample children's understanding of conventional time terms such as "after," "minute" and "year." His second test was administered to children in grades four, five and six and was designed to follow up the concepts sampled in the first test and to look at a wider variety of time concepts, such as those involved in the use of indefinite time expressions, time lines and chronological sequence.

By the time children reach grade six, Friedman concluded, they have a satisfactory understanding of conventional time terms but an unsatisfactory understanding of chronological order and specific time words and dates. It

¹⁶Friedman I, op. cit., pp.337-42 .

was found that time lines were understood by only a very few pupils; in fact, by only 27% of pupils at the grade six level. Factors of sex, I.Q., and socio-economic status were not found to contribute significantly to the findings of the overall study.

In a second study¹⁷ of children's time concepts Friedman involved six hundred and sixty-seven pupils in grades seven to twelve as well as one hundred and ninety-four adults. The measuring instrument used consisted of four tests. Test one dealt with the meaning of dates and words descriptive of time such as "ancient". Test two incorporated questions about groups of items to be remembered in chronological sequence thus testing perspective on historical and current events. Tests three and four consisted of questions on space-time relationships with specific reference to time lines. Adults did not do the latter two tests.

Upon analysis of the responses, Friedman concluded that the results of his time line test demonstrated that pupils can understand a simple historical time line as early as the eighth grade level. However, a line involving complexities in dating should be used guardedly he claims.

As one of the predictor variables, intelligence quotient was found to have a marked, but not high, positive

¹⁷Friedman II, op. cit., pp. 233-38.

correlation with performance on the time comprehension test. On the other hand, sex and socio-economic status were not significantly related to the understandings covered in the tests.

In an extension of the above study, Friedman and Marti¹⁸ investigated the effects of training and increased age on maturation of the time sense. Friedman's four-part 1944 test of time comprehension¹⁹ was used as the pre- and post-test instrument.

Three grade ten groups were set up and maintained over one semester's time. Group one, a world history group of seventy-two pupils, received instruction on time concepts through teaching in class and various exercises such as making time lines and studying chronological sequences and historical time concepts. Group two, consisting of eighty-eight pupils who were also taking world history, received no instruction on time concepts. Group three, a group of two hundred and thirteen pupils who took no world history, received no instruction on time concepts either.

Friedman and Marti's study showed evidence that over a semester's duration the ability to comprehend time can be developed through systematic instruction at the junior high and high school levels. Further, they concluded that,

¹⁸K.C. Friedman and V.A. Marti, "A Time Comprehension Test," Journal of Educational Research, XXXIX (September, 1945), 62-68.

¹⁹Friedman II, op.cit., pp. 233-38.

although the time sense matures with age and experience, social studies teachers might contribute greatly to its acceleration through systematic instruction in such activities as making time lines, putting events in chronological sequence, and using historical time concepts.

Maturation of the time sense in children of from age eighteen months to eight years was studied further by Ames in 1946.²⁰ The study was based on a sample of ninety-five subjects having high or superior intelligence. Children from eighteen to forty-eight months of age were observed through a one-way-vision screen as they participated in free or directed play and their verbalizations were recorded and analyzed over a period of two years. In addition, the children were asked a series of questions designed to test time concepts. Children in the five to eight year age group were asked the questions only and were not observed in the same way as the younger children.

Ames concluded that within any one level of age and intelligence there are marked individual differences in children's concepts of time. But despite these differences, understanding of time concepts seems to develop in a relatively set pattern and at approximately the same time for all children. This finding led Ames to hypothesize that

²⁰L.B. Ames, "The Development of the Sense of Time in the Young Child," Pedagogical Seminary and Journal of Genetic Psychology, LXVIII (April, 1946), 97-125.

that maturation is a vital factor in the development of time concepts.

In 1947, Bradley²¹ carried out a replication of the Oakden and Sturt study²² using samples of from fifty-seven to one hundred and twenty-seven subjects. From the study Bradley concluded that by age seven the child has a relatively sound knowledge of time terms. He also found that, although nine-year-olds can comprehend a long period of years, understanding of measures of duration does not reach satisfactory levels until about age twelve.

Generally the other findings support those of Oakden and Sturt except that the sudden improvement in time knowledge at about age eleven as reported by the earlier study was not observed. Thus, Bradley saw the acquiring of time knowledge as a gradual, continuous process. A satisfactory reason for the difference from Oakden and Sturt was not advanced by Bradley although it has been suggested that differences in samples might account for the discrepancy.²³

Another study in 1952 by Callahan concentrated on rates of development of various time concepts such as order and duration.²⁴ One hundred and ninety-two fifth graders and

²¹Bradley, op. cit., pp. 67-78.

²²Oakden and Sturt, op. cit., pp. 309-36.

²³Newman, op. cit., p. 13.

²⁴Callahan in Chase, "American History," op. cit., pp. 337-38.

two hundred sixth graders attempted to place in chronological order five items in each of twenty-four questions.

Callahan saw from the results the anticipated better understanding of the sixth graders as well as significant differences between individuals at each grade level. In addition he found that the boys appeared more capable of arranging events in chronological order than girls. Finally, when events were related to a present rather than a past event children at both grade levels were better able to put events in order.

Farrel, in a 1953 study,²⁵ tried to test understanding of time relations in five, six and seven-year-old children of high I.Q. In addition she attempted to discover any relationship between time understanding and chronological age, mental age, and sex. The subjects were seventy-five "gifted" children.

Her results seem to indicate a growth of time understanding between the five- and six-year-old levels but not between the six- and seven-year-old levels. Such a lack of growth might be a plateau or perhaps the questions were too easy for the seven-year-olds to do better than the six-year-olds.

Sex differences were found to be insignificant as predictors of success. Also, time questions which involved

²⁵Farrell, op. cit., pp. 587-94.

the personal and immediate were better and more correctly answered at a lower chronological and mental age than questions which involved the non-personal and remote.

Arnsdorf's 1961 investigation²⁶ of the teaching of chronology in the sixth grade attempted to assess the effects of a seven week period of instruction utilizing the vocabulary of chronology in basic texts. The subjects, five hundred and sixty-three sixth graders, were divided into two groups. Both groups were taught the same social studies unit from the same text for a period of seven weeks. However, one group was given specific instruction about the time terms encountered in the unit. After seven weeks an original battery of six tests to measure understanding of time concepts was administered.

The results showed that, according to the tests, pupils in the experimental groups had a significantly greater understanding of time concepts than those in the control group. However, it should be mentioned that all subjects were from self-contained classrooms and there were no controls for previous experiences except random selection of classrooms for the control and experimental groups.

An additional study in 1961, by McAuley, attempted to determine what understandings second grade children have of

²⁶V.E. Arnsdorf, "An Investigation of the Teaching of Chronology in the Sixth Grade," Journal of Experimental Education, XXIX (March, 1961), 307-13.

time relationships.²⁷ From an oral, individually administered test McAuley found that the subjects had difficulty associating the past with the present in the immediate environment but they were capable of understanding periods of time particularly when pivoted on events.

Most subjects had some understanding of past social reality if it was not related directly to the immediate concerns or environment of the subjects. Also, most showed little understanding of continuity on a personal basis as between mother and child but showed a better understanding of calendar continuity. For example, the child is better able to understand time relationships between pioneers and Indians better than between his mother and himself.

As should be pointed out this study seems to contradict the findings of an earlier study by Farrel. Farrel found that time questions which involved the personal or immediate were better and more correctly answered at a lower chronological and mental age than questions which involved the non-personal and remote.²⁸

Interpretations of indefinite expressions of time such as "in colonial days" and "middle ages," studied earlier

²⁷J.D. McAuley, "What Understanding do Second Grade Children Have to Time Relationships?" Journal of Educational Research, LIV (1961), 312-14.

²⁸Farrell, op. cit., p. 94.

by Kelty²⁹ were investigated by Gill in 1962.³⁰ Using a questionnaire containing eighteen indefinite expressions of time which the subjects were to give definite numerical quantification or dates for, Gill tested two hundred and sixty subjects in four nearly-equal sized groups based on college, high school, eighth grade, and fifth grade levels.

He found that words like "century" and "decade" were not clearly understood by many and terms such as "ancient time" or "modern times" had no precise meaning for many, particularly in the lower grades. Generally, higher grade levels, especially college, had a superior grasp of the meanings thus tending to confirm studies where time-sense and maturity appear closely related.

Another study of the benefits from systematic instruction about time concepts was done by Dobbs in 1963.³¹ After setting up control and experimental groups with appropriate pre-tests the experimental group received systematic instruction for six weeks by means of trade books written on the topic time as well as in arithmetic and science.

Results of a post-test indicated that the instruction

²⁹Kelty, op. cit., pp. 607-18.

³⁰Gill, op. cit., pp. 454-56.

³¹E.L. Dobbs, "A Study of the Sequential Development of Time Sense and Chronology in the Elementary School," (unpublished Ed. D. thesis, University of Kansas, 1963), pp. 73-87.

was beneficial. However, it was noted that in the experimental group the boys had a greater ability to estimate intervals of time than the girls while the latter had a more extensive vocabulary of time words.³²

Finally, a 1967 study by Newman attempted to investigate the factors which influence the elementary school child's understanding of time duration and the stages by which such understanding is acquired.³³ A battery of three sub-tests on understanding of verbal and graphic comparisons of duration as well as on ability to use measures of duration was administered to one hundred and ninety-two children in grades one to six. There were sixteen boys and sixteen girls at each grade level.

Among the findings,³⁴ intelligence, chronological age, and grade placement were all significant predictors of ability to understand duration. Socio-economic status was found to be a significant predictor of ability to understand verbal comparisons and measures of duration but not graphic comparisons of duration. Sex was not significantly related to the ability to understand duration. Comparisons of duration were generally understood by children at the grade

³²Ibid., pp. 83-85.

³³Newman, op. cit., Abstract and pp. 96-104.

³⁴Ibid., pp. 96-104.

three level while measures of duration were first satisfactorily comprehended at the grade four level. At the grade six level there was a noticeable improvement in the ability to understand comparisons and measures of duration.

RELATING THE HYPOTHESES TO THE LITERATURE

The four hypotheses of the present study were concerned with the effects of four variables: intelligence, reading level, age and sex. Reasons why these variables were chosen and why they merited inclusion in hypotheses will be given in the discussion of the literature following each hypothesis.

Hypothesis one states:

In terms of pupil achievement on a chronological relationships test the effect of a historical narrative with a time line is superior to the effect of the narrative alone at the high intelligence level.³⁵

It should be noted that the high intelligence group with the illustrative time line is expected to benefit. Support for the assertion that such an intelligence group will benefit from the time line comes from the study of time line effectiveness carried out by Davis, Hicks and Bowers³⁶

³⁵See Chapter 3 for discussion of intelligence levels pp. 62-63.

³⁶Davis et al, op. cit., pp. 22-25

as well as from Vernon's study of the usefulness of graphs as illustrations.³⁷

It should also be noted that Friedman³⁸ and Newman³⁹ have found that intelligence is a predictor of ability to perform on a time comprehension test and to understand duration, respectively. It seems likely that ability to comprehend time could aid one in comprehending time lines. Similarly, understanding of duration could aid one in understanding time lines since time lines are spatial representations of events and the durations of time between them.

It may be noted that the finding by Davis, Hicks, and Bowers that high intelligence students benefit from an illustrative time line was contradicted by Davis' later study⁴⁰ which showed no significant difference between time line and non-time line groups regardless of intelligence level. However, as pointed out in an earlier criticism of both studies,⁴¹ intelligence may not have had much opportunity to operate since the studies were so dependent on memory. Intelligence may have had more opportunity to operate in the first study since it incorporated only a 700

³⁷Vernon, op. cit., pp. 180-87.

³⁸Friedman II, op. cit., pp. 237-38.

³⁹Newman, op. cit., p. 96.

⁴⁰Davis, op. cit., p. 752

⁴¹See pp.14-15.

word narrative to be studied with the time line as opposed to a 1150 word narrative with the time line in the second study. Furthermore, the students in the first study were grade eleven and twelve history students who would be more mature intellectually and possible less reliant on memory when compared to the junior high school students used in the second study by Davis.

Hypothesis two states:

In terms of pupil achievement on a chronological relationships test, the effect of a historical narrative with a time line is the same as the effect of the narrative alone at below average, average, or above average reading comprehension levels.⁴²

Thus, the hypothesis indicates that reading level is expected to affect both treatment groups equally.

The reason why reading level has been chosen as an independent variable is that reading ability, as indicated by reading comprehension level, will affect comprehension of the written narrative and consequently the scores on the chronological relationships test. However, according to several studies of the influence of illustrations on reading comprehension it is not possible to infer that illustrations, such as time lines, affect reading comprehension.⁴³ Thus, there

⁴²See Chapter 3 for discussion of these levels, pp.56-57.

⁴³Burdick, op. cit., p. 37; Vernon, op. cit., p. 186; Miller, op. cit., pp. 181-82.

is no basis from which to predict which of the three reading comprehension levels will benefit most from the illustrative time line, if any do. Consequently, reading comprehension level is stated as having the same affect in both treatment groups since there is no evidence to show otherwise.

Hypothesis three states:

In terms of pupil achievement on a chronological relationships test the effect of a historical narrative with a time line increases equally with the effect of the narrative alone as chronological age increases.

From this hypothesis it can be seen that chronological age is expected to affect both treatment groups equally.

A number of studies have shown that chronological age is a predictor of ability to deal with time and chronology. For example, Oakden and Sturt pointed out that knowledge of conventional time grows from at least age four on and that by age eleven the child has some of the concepts necessary for the study of history.⁴⁴ Later studies by people such as Kelty, Harrison, Bradley, Farrel, Callahan, Gill, and Newman⁴⁵ are among those showing that various aspects of the understanding of time and chronology improve at higher

⁴⁴Oakden and Sturt, op. cit., pp. 334-36.

⁴⁵Kelty, op. cit., pp. 615-18; Harrison, op. cit., pp. 513-14; Bradley, op. cit., pp. 77-78; Callahan in Chase, "American History," op. cit., pp. 337-38; Farrel, op. cit., pp. 592-94; Gill, op. cit., pp. 455-56; Newman, op. cit., pp. 96-97.

chronological age levels.

Such research provides no evidence to indicate whether an illustrative time line is more effective for older children. At high chronological ages it seems likely that the pupils will be better able to comprehend the order of dates and the durations of time described in the narrative insofar as their understanding of time and chronology improves with age. However, such improvement does not necessarily mean that those with the time line will put it to more effective use and thus score higher on the criterion test⁴⁶ than those without the time line. Indeed, it is not known whether pupils with the time line will use it. In addition, there is no way to know if the ability to use the time line will improve with age or if the age span in months within a single grade, as in the present study, is sufficient to make any difference in such ability. The pertinent question is: Will the effectiveness of the time line increase as chronological age increases?

Since there is no strong basis for predicting increased effectiveness with increased chronological age, hypothesis two states that if the pupils in each of the two treatment groups are compared they will improve equally in performance as chronological age increases. Finally, such a prediction of no-difference receives more support from Davis'

⁴⁶Chronological Relationships Test.

study⁴⁷ of time line effectiveness where he found no significant difference in scores for time line and non-time line groups when controlling for grade level, insofar as different grade levels can be equated with different age levels.

Hypotheses four states:

In terms of pupil achievement on a chronological relationships test the effect of a historical narrative with a time line is the same as the effect of the narrative alone for boys or for girls.

It can be shown that there is some contradictory evidence on whether the sex of the child appears to influence the ability to comprehend concepts of time and chronology.

In Friedman's first investigation of elementary school children's time concepts⁴⁸ he found that sex did not contribute significantly to the findings of the overall study. He came to the same conclusion after his second such investigation, at the junior high, high school and adult levels.⁴⁹ Similarly, Farrel, when testing the understanding of time relations in the four-, five- and six-year-olds⁵⁰ found that sex differences were insignificant as predictors of success. Finally, Newman concluded that sex was not significantly

⁴⁷Davis, op. cit., p. 752.

⁴⁸Friedman I, op. cit., pp. 337-42.

⁴⁹Friedman II, op. cit., pp. 237-38.

⁵⁰Farrel, op. cit., pp. 587-94.

related to the ability to understand duration, at the grades on through six levels.⁵¹

There is some contradictory evidence, however. First Callahan found that boys appeared more capable of arranging events in chronological order than girls.⁵² Furthermore, Dobbs found that after six weeks of instruction about time concepts the boys in her experimental group had a greater ability to estimate intervals of time than the girls, However, girls had a more extensive vocabulary of time words.⁵³

From the findings of Friedman, Farrel, and Newman it appears that generally there is no difference between the sexes in ability to comprehend time concepts. But, the findings of Callahan and Dobbs suggest that boys and girls may excell in differing aspects of time comprehension.

If boys can arrange events in chronological order or can learn to estimate intervals of time better than girls they might be able to read or use a time line, showing a chronological sequence and intervals of time, better than girls. Such an effect could result in boys with the illustrative time line doing better than boys without the time line who have to depend solely on the written narrative.

⁵¹Newman, op. cit., p. 96.

⁵²Callahan in Chase, "American History," op. cit., pp. 337-38.

⁵³Dobbs, op. cit., p. 85.

On the other hand, if girls tend to develop a more extensive vocabulary of time words perhaps they will tend to exercise their vocabulary, thus depending on the verbal narrative and neglecting the time line. Such an effect could result in the girls performing similarly with or without the illustrative time line.

Such speculation about the abilities of boys and girls helps to point out some possible ways in which sex might affect scores on the criterion test. Although there is not enough evidence to suggest whether boys or girls will benefit most from the time line it appears that the inclusion of hypothesis four is justified since sex might be associated with specific abilities required for reading a narrative or a time line.

In addition to intelligence, reading level, age and sex there are at least two other factors which might affect the usefulness of a time line for an individual. These are training in how to use a time line, and socio-economic status.

For several reasons it appears difficult if not impossible to determine the amount of training in the use of time lines which pupils have had by the time they reach grade six. Presumably some teachers would stress instruction and practice in the use of time lines more than others. Apparently, such training might affect a pupil's ability to use a time line since training appears to increase under-

standing of some aspects of time and chronology.⁵⁴

Since pupils usually have different teachers each year in elementary school it is difficult to get an accurate ex post facto picture of the past training of each individual pupil. Such an undertaking was considered beyond the scope of the present study since it would be a major piece of work to attempt to piece together a picture of each child's past training in time line usage. Some teachers would have moved to different schools or areas and the accuracy of their memories as well as those of the pupils would be questionable. However, to gain some information about training in time line usage this investigator gave each participating grade six teacher a questionnaire on teaching of chronology over the current year.⁵⁵

To further minimize the need to assess the effect of training on classes which might have received more training on time line usage the assigning of pupils to the two treatments has been randomized within each class for the present study.⁵⁶ Thus, it is unlikely that pupils who have had training in time line usage might have formed a preponderance of the subjects on treatment one, the narrative plus time

⁵⁴Friedman and Marti, op. cit., pp. 67-68; Arnsdorf, op. cit., p. 313; Dobbs, op. cit., pp. 75-78.

⁵⁵See Appendix E.

⁵⁶See discussion of randomization procedures in Chapter 3, p.72.

line, Consequently, there is little likelihood that because of more training in the use of time lines the subjects with the illustrative time line plus narrative will excell over the subjects with the narrative only. Both groups of subjects, through randomization, "the most adequate all-purpose assurance of lack of initial biases between groups,"⁵⁷ should have approximately the same amount of training in time line usage. Any differential effects between the groups should be attributable to the nature of the treatments.

A final factor which might affect the present study is socio-economic status. However, there is conflicting evidence on the effect of such status, as well as procedural difficulties in determining the socio-economic status of pupils.

In Friedman's first and second studies on children's and adult's time concepts he found no significant difference between various socio-economic classes.⁵⁸ In both studies he classified the pupils according to the occupations of their parents, as rated by the Minnesota Occupational Scale.⁵⁹

⁵⁷D.T. Campbell and J.C. Stanley, Experimental and Quasi-Experimental Designs for Research (Chicago: Rand McNalley, 1963), p. 25. For further discussion, Ibid., p. 2 and see also discussion in F.N. Kerlinger, Foundations of Behavioral Research (New York: Holt, Rinehart and Winston, 1964), pp. 56-63.

⁵⁸Friedman I, op. cit., pp. 240-41; Friedman II, op. cit., p. 337.

⁵⁹Friedman I, op. cit., p. 233; Friedman II, op. cit., p. 337.

Unlike Friedman, Newman found a significant relationship between socio-economic status and the ability to understand comparisons and measures of duration.⁶⁰ Also, he determined socio-economic status by drawing half of this sample from a school judged to be in a relative low socio-economic area and the other half from a school judged to be in a relatively high socio-economic area. The schools were selected on the basis of advice from personnel attached to the school board.⁶¹

In the present study this investigator has decided not to attempt to examine the relationship between socio-economic status and the effectiveness of illustrative time lines. There are several reasons for the decision.

First, due to the apparent heterogeneous distribution of socio-economic status within each school in Grande Prairie, where the present study has been carried out, a procedure similar to Newman's could not be used. Second, if one is to use some other procedure in order to determine the socio-economic status of each pupil, rather than each school, then suitable criteria for determining socio-economic status must be set up.

One of the better known scales for determining socio-economic position is the Blishen Scale which utilizes a

⁶⁰Newman, op. cit., p. 96.

⁶¹Newman, op. cit., pp. 30-31.

ranking system based on occupation.⁶² Other approaches which could be used to determine socio-economic status are the self-classification approach, the objective approach, or the reputational approach.⁶³ However, the latter three approaches were not viable because of lack of means to get necessary data from the parents in the community.

Third, an occupational scale of socio-economic classification such as that of Blishen⁶⁴ may be useful to sociologists in studying mass groups. However, this investigator feels that there are sufficient variations in individual characteristics within occupational groups to make any conclusions drawn from such grouping too general to be applicable for many individual pupils. For example, in the present study one might have come to the hypothetical conclusion that lower socio-economic class pupils did not benefit as much from the illustrative time line as middle or upper socio-economic class pupils.

The question is: If one of the conclusions of the present study was a statement about how effectively a time line was utilized by a socio-economic class based upon the

⁶²B.R. Blishen, "The Construction and Use of an Occupational Class Scale," Canadian Journal of Economics and Political Science, XXIV (Nov., 1958), pp. 519-31.

⁶³G.A. Lundberg, C.C. Schrag, and O.N. Larsen, Sociology (rev.ed.; New York: Harper and Row, 1963) pp.341-44.

⁶⁴Blishen, op. cit., pp. 519-31.

occupation of the pupils' parents, then how useful would such knowledge be?

Whether a particular child used the time line would have been determined by more specific characteristics of the child, such as intelligence or reading ability, which might be distributed differently in different socio-economic classes. However, as discussed in the current chapter, the present study has included such specific characteristics as independent variables, thus eliminating the necessity of including broad socio-economic classes as independent variables.

Another variable, mental age, which is related to maturation, has not been included in the present hypotheses either. The reason why it had been omitted is that statistics on mental age are usually not readily available to or used by teachers in deciding whether teaching aids such as time lines are appropriate for specific pupils. Thus chronological age, which is readily obtained by teachers, as well as I.Q., which is usually a part of cumulative records, are included as independent variables in hypotheses one and three.⁶⁵ However, the present investigator also felt that if mental age rather than chronological age later appeared to be a more useful indicator of maturation, as in Harrison's findings⁶⁶, then an

⁶⁵See p. 31 and p. 34, respectively.

⁶⁶Harrison, op. cit., pp. 513-14.

a posteriori hypothesis on the possible effect of mental age in the present study could be stated and tested.

Possibly, further hypotheses about the effect of arithmetic ability, vocabulary or other variables could have been elaborated for this study. However, due to the need to gather more data in order to test more hypotheses, the investigation of additional variables appeared to be beyond the scope of the present study. Also, when tested, the four hypotheses which have been elaborated for the present study should provide an indication of the answer to the central problem which in question form was: Is the effect of an illustrative time line with a historical narrative superior to the historical narrative alone in terms of pupil understanding of the chronological relationships in the narrative?

Within the context of the research literature of this chapter, four hypotheses about the effectiveness of illustrative time lines have been stated and related to that literature. Included as variables in those hypotheses are intelligence, reading ability, chronological age, and sex. In the following chapter the methods of gathering data relevant to the support or non-support of each hypothesis have been elaborated.

Chapter 3

METHODOLOGY

Within this chapter, first the design of the experiment is discussed then the variables are reviewed. Next, descriptions of the sample, the instruments, and the data-gathering procedures are provided. Finally, the null hypotheses tested are listed and the statistical procedures used are described.

DESIGN OF THE EXPERIMENT

This study used the posttest-only control group design.¹ Schematically the design is:

R	X	O ₁
R		O ₂

¹Campbell and Stanley, op. cit., pp. 25-26; W. Wiersma, Research Methods in Education (New York: J.B. Lippincott, 1969), pp. 228-30; or the experimental group-control group: randomized subjects design in F.N. Kerlinger, Foundations of Behavioral Research (Holt, Rineheart, and Winston: New York, 1964), pp. 303-04.

where R refers to the fact that students are randomly assigned to the experimental and control groups, X refers to the experimental treatment, and O_1 and O_2 refer to the criterion measure of the experimental and control groups, respectively. For a number of reasons this design appeared suitable for the present study.

First, pupils within a classroom can be prevented from realizing that they are on different treatment materials since randomization can be handled by a mixed ordering of materials for distribution.² For this study two sets of similar-looking treatment materials³ were distributed in each class with each pupil receiving one or the other set.

Second, the posttest-only control group design does not require pretesting.⁴ In the present study a pretest was not available and did not appear to be necessary since experimental and control groups could be compared by their performance on the criterion test alone.

Third, when used in a multivariate case, the design can test several hypotheses at one time.⁵ The present study involved multiple variables which were intelligence, reading comprehension, age, and sex. Within the one design the four

²Campbell and Stanley, op. cit., p. 26.

³See Appendix A.

⁴Wiersma, op. cit., p. 228.

⁵Kerlinger, op. cit., p. 303.

directive hypotheses⁶ incorporating the four variables were tested.

Fourth, the design appeared particularly effective from a research point of view for two reasons:⁷

(a) The control group provides a comparison for the experimental group so that it can be determined if the experimental treatment has any effect. Accordingly, within the present study the experimental or time line plus narrative group has been compared with the control or narrative only group according to respective pupil performances on the criterion test.

(b) The design has the virtue of randomization.⁸ Within the present study in order to make the experimental group-control group idea valid, it has been necessary for the experimenter to have some assurance that his groups were approximately (statistically) equal on any variables possibly related to the dependent variable, time line comprehension. Neither matching nor intuitive or experienced judgements could accomplish such equalization.⁹ Thus, the pupils were assigned randomly to the experimental and control groups.¹⁰

Fifth, the external validity of the posttest-only

⁶See Chapter 1, pp. 7-8.

⁷Kerlinger, op. cit., p. 304.

⁸Ibid.

⁹Ibid.; see also discussion of matching in Ibid., p.286.

¹⁰See discussion in procedures, p. 72.

control group design is high, provided that the experimental and non-experimental settings are similar enough for generalizing from one to the other. If generalizing can be done then the random assignment of subjects is assumed to have eliminated the possibility of selection biases in the experimental group interfering with generalization.¹¹ Thus, if others wish to generalize to populations other than the sample used in this study they should be able to do so confidently to the extent that those populations can be shown to be like the sample used in the present study.

In contrast to the above-mentioned strengths the posttest-only control group design has at least two weaknesses which could threaten internal validity. They are differential mortality and the possible entrance of relevant but non-experimental factors into the experiment.¹²

Differential mortality means that the dropouts in one group have characteristics different from those of the other group and that these characteristics may be relevant to the experimental and dependent variables. However, in the present study differential mortality could not take place since the experimental and control treatments as well as the criterion test were administered within one class period. Thus, no subjects were lost during the experiment.

¹¹Wiersma, op. cit., p. 229.

¹²Ibid.

The entrance of relevant but non-experimental factors into the experiment could have happened in the present study if, for example, the experimental group performed better on the criterion test because they felt unique, since they had a time line which the control group did not have. To guard against such a "Hawthorn Effect" or other effects due to other relevant but non-experimental factors this study has been carefully planned as shown in the procedures¹³ and care has been taken to ensure that pupils on experimental and control treatments were treated the same.

By way of summary, the following five strenths of the posttest-only control group design appear to have been of benefit in this study:

(1) Randomization has been handled by the mixed ordering of materials for distribution.

(2) No pretesting has been required.

(3) Several hypotheses have been tested by the one design.

(4) The design appeared effective from a research point of view by;

(a) having the comparability of a control group.

(b) validating the experimental group-control group idea through randomization.

(5) External validity appeared to be high.

¹³See pp. 70-73.

One weakness of the design has been the possible unseen entrance of relevant but non-experimental factors into the experiment. However, the procedures attempted to prevent such factors from entering.¹⁴

THE DEPENDENT AND INDEPENDENT VARIABLES

Understanding of the chronological relationships described in the narrative and shown on the time line, as indicated by score on the Chronological Relationships Test, was the dependent variable. The independent variables were presentation method, intelligence, reading comprehension, age, and sex.

Presentation method had two levels consisting of (a) a narrative with an illustrative time line for the experimental group¹⁵ and (b) the same narrative without the illustrative time line for the control group.¹⁶

Intelligence had four levels which were (a) low (b) average (c) high average and (d) high.¹⁷

Reading comprehension had three levels which were

¹⁴Ibid.

¹⁵See Appendix A, Part 1.

¹⁶See Appendix A, Part 2.

¹⁷See how these were determined on pp. 62-63.

(b) average and (c) high.¹⁸

Age was a discrete variable consisting of the pupils' ages in months.

Finally, sex had two levels, being either male or female.

SAMPLE

All of the classes of grade six pupils in the Grande Prairie Public School District were used, making a total sample of 214 pupils. The size of the sample was adequate for dividing into four sufficiently large groups in order to accommodate analysis of differences between means of the I.Q. variable which had the maximum number of four levels. As it turned out, the smallest group was the low I.Q. sample which consisted of 24 subjects, 10 on the experimental treatment and 14 on the control treatment, which was not a sufficiently large number to confidently perform a t-test for significance on the difference between mean scores in the two treatment groups.¹⁹ All other groups had over the 30 subjects suggested as adequate for performing the t-test.²⁰

¹⁸See how these were determined on pp. 56-57.

¹⁹Wiersma, op. cit., p. 75; thus, the non-parametric median test was performed to check results of the t-test on the sample of 24.

²⁰Wiersma, op. cit., p. 75.

VALIDITY AND RELIABILITY OF THE TEST INSTRUMENTS USED

In this section the validity and reliability of the three test instruments used in the present study will be assessed. For the purposes of this study the term "validity" has been defined as the "accuracy with which the test measures what it is intended to measure."²¹ "Reliability" by contrast has been defined as the "accuracy with which a test measures whatever it does measure."²²

Instruments Used

Three test instruments were used in this study. One instrument consisted of the reading comprehension test from the Canadian Tests of Basic Skills, Form 1,²³ grade six level, which was used to determine reading comprehension level. A second instrument was the California Short-Form Test of Mental Maturity, Level 2,²⁴ which was used to deter-

²¹R.L. Ebel, Measuring Educational Achievement (Englewood Cliffs: Prentice-Hall, 1965), pp. 388-89.

²²Ibid.

²³E.M. King, ed., in cooperation with E.F. Lindquist and A.N. Hieronymus, Canadian Tests of Basic Skills, Form 2 (Toronto: Thomas Nelson and Sons, 1967).

²⁴E.T. Sullivan, W.W. Clark and E.W. Tiegs, California Short-Form Test of Mental Maturity (Monterey, Calif.: California Test Bureau, 1963).

mine intelligence quotient. The third instrument was called the Chronological Relationships Test²⁵ and was used to measure pupil understanding of the material incorporated into the historical narrative as well as on the time line.

Each test in turn will now be described, validity and reliability assessed, and use in the present study explained.

Reading comprehension test. In the present study the test used to determine reading level was the reading comprehension test from the Canadian Tests of Basic Skills, Form 1²⁶ or CTBS-1. The latter is a battery of tests designed to measure vocabulary, reading comprehension, language, work-study skills, and arithmetic skills.

1. Description. The CTBS-1 grade six reading comprehension test consisted of 76 multiple-choice items which were to be administered in 55 minutes.²⁷ Items on the grade six test were part of a total test, successive parts of which were intended for use anywhere from grade three to grade eight. The grade six test, as with the other tests, overlapped on some items in common with the tests intended

²⁵See Appendix B.

²⁶King, op. cit.

²⁷King, Manual For Administrators, Supervisors and Counsellors, (Hereinafter referred to as Manual I), op. cit., pp. 4-5.

for grades above and below grade six.

Selections upon which the test questions were based varied in length from two to seven paragraphs. They included selections dealing with science, anthropology, history, a literary anecdote, and biography.²⁸ Such selections had come from various sources such as newspapers, magazines, literary works and textbooks.²⁹

Questions on the test attempted to assess the following skills:³⁰

Details - ability to recognize and understand stated or implied factual details and relationships. These included such things as specific facts, implied facts, and deduced meanings.

Purpose - skill in discerning the purpose or main idea of a paragraph or selection.

Organization - ability to organize ideas through the recognition of common elements, parallel topics or proper time sequence.

Evaluation - skill in evaluating what has been read by such activities as generalizing from a selection, recognizing a writer's viewpoint or a mood and recognizing outstanding qualities of style.

²⁸from examination of the test selections by the present writer.

²⁹King, Manual I, op. cit., p. 30.

³⁰Ibid.

Scores on the CTBS-1 reading comprehension test were able to be translated into grade-equivalent scores or percentile scores. The latter were based on beginning-of-the-year, and end-of-the-year norms and a separate table was given for each of grades three through eight.³¹

2. Assessment of validity and reliability. The first 60 out of the total 76 items on the CTBS-1 grade six reading comprehension test were exactly the same as the first 60 out of the total 76 items on the equivalent test in the Iowa Tests of Basic Skills, Form 3.³² Since the rest of the grade six reading comprehension items on the CTBS-1 were developed on the same principles³³ as those of the four forms of the ITBS³⁴ it appeared that CTBS-1 was measuring the same abilities as the ITBS batteries. From these facts it appeared that criticism of the validity of the ITBS batteries would be applicable to the CTBS-1 in the absense of critical material on the relatively new CTBS-1 battery.

Considerable evidence of direct validity in the form of agreement by several independent reviewers that the ITBS, and thus the CTBS-1 test items were measuring reading compre-

³¹King, Teacher's Manual, Form 1, (Hereinafter referred to as Manual II), op. cit., pp. 69-75.

³²E.F. Lindquist and A.N. Hieronymus, Iowa Tests of Basic Skills (Boston: Houghton Mifflin, 1964).

³³King, Manual I, op. cit., p. 6.

³⁴Copyrighted in 1955, 1956 and 1964 at the State University of Iowa.

hension has been given.³⁵ It appeared that a good measure of effort had been expended by the CTBS-1 test constructors in attempting to set up and abide by valid principles for scaling and standardization of the CTBS-1.³⁶ Furthermore, the ITBS reading comprehension subtest appeared to have some predictive validity.³⁷ Finally, a reliability of .92 for the CTBS-138 reading comprehension test seemed to be satisfactory.³⁹

In view of the above summary of evidence the CTBS-1 reading comprehension test has been considered to have adequate validity as a test of reading comprehension for the purposes of the present study. In this study test scores have been used as described in the next section.

3. Use in the present study. Within the present study reading comprehension as an independent variable had three levels. These levels were determined by the following cut-off points in terms of grade-equivalent scores.⁴⁰

³⁵in O.K. Buros, Fifth Mental Measurements Yearbook, (Highland Park, N.J.: Gryphon Press, 1959), pp. 30-37; for a discussion of Buro's yearbooks see Hoffman, op. cit., pp.38-40.

³⁶King, Manual I, op. cit., p. 6 and pp. 50-51.

³⁷Ibid., pp. 38-45.

³⁸Ibid., p. 38 and p. 50.

³⁹N.E. Gronlund, Constructing Achievement Tests (Englewood Cliffs: Prentice-Hall, 1968), p. 96.

⁴⁰as determined by grade-equivalent conversion tables where the first digit stands for school year and the second for month in King, Manual II, op. cit., p. 92.

Level	Grade Equivalent Score
low	5.9-
average	6.0 to 6.9
high	7.0+

Using such cut-off points to group the sample scores into three reading comprehension ability groups seemed justified for at least two reasons. First, the present study was conducted on a sample of grade six pupils so that it seemed appropriate to select those in the grade six interval to represent the grade six reading comprehension level. Similarly, those scoring less than the interval seemed to be an appropriate group to represent pupils with ability below the grade six level and those with ability at the grade seven level and above seemed to be an appropriate group to represent the high reading level.

Dividing the pupils' scores into groups on the basis of reading level was somewhat arbitrary however⁴¹ and the method of determining the groups should be kept in mind. As a result of the division, the low, middle, and high groups contained the following numbers of pupils, respectively: forty-three, sixty-eight, and ninety-five.

California Short-Form Test of Mental Maturity. For

⁴¹None of the previous studies of time lines had used reading comprehension as a variable. Thus, no model for determining levels was available for the present study.

determining intelligence level within the present study the California Short-Form Test of Mental Maturity,⁴² Level 2, was used. The Short-Form is a one period adaptation of the California Test of Mental Maturity.⁴³

In the immediately following section information on the suitability of the Short-Form has been given. The test has been described first then the validity and reliability have been analyzed and appraised. Finally, the employment of the test in the present study has been explained.

1. Description. To cover the grade and age range from preschool to adult, the Short-Form appears at eight articulated levels. These levels, followed by the grades they span, are Level 0 (K-L1), Level 1 (H1-3), Level 2 (4-6), Level 2H (6-7), Level 3 (7-9), Level 4 (9-12), and Level 5 (12-Adult).⁴⁴

At each level, the rate and scope of mental development are measured in terms of four statistically-derived

⁴²E.T. Sullivan, W.W. Clark, and E.W. Tiegs, California Short-Form Test of Mental Maturity, 1963 Revision, (Hereinafter referred to as Short-Form), (Monterey, Calif.: California Test Bureau, 1963).

⁴³E.T. Sullivan, W.W. Clark, and E.W. Tiegs, California Test of Mental Maturity, 1963 Revision, (Hereinafter referred to as Long-Form), (Monterey, Calif.: California Test Bureau, 1963).

⁴⁴Sullivan et al, Examiners Manual, Short-Form, (Hereinafter referred to as Manual), op. cit., p. 8.

factors:⁴⁵ Logical Reasoning, Numerical Reasoning, Verbal Concepts, and Memory. Within these factors, the seven test units are grouped into two sections, Language and Non-Language, that differentiate in general between responses to stimuli that are essentially nonverbal or pictorial.⁴⁶

In regard to the length of the test; It requires 43 minutes to administer, not including time for reading directions, and contains 120 multiple-choice items.⁴⁷ The Non-Language and Language Sections each contain 60 items. All items are arranged in ascending order of difficulty in each test unit.⁴⁸

Interpretation of an examinee's performance on the Short-Form is facilitated by the individual profile sheet. This sheet permits the making of a graphic record of factor, section and total scores in terms of standard scores, stanines, or percentile ranks. Furthermore, the Language and Non-Language sections' scores as well as the total scores may each be converted to I.Q. scores or mental ages.⁴⁹

⁴⁵derived using the Thurstone centroid method mentioned in Sullivan et al, Manual II, op. cit., p. 7; see explanation in Lindvall, op. cit., pp. 151-52.

⁴⁶Sullivan et al, Manual, op. cit., p. 5.

⁴⁷Ibid., p. 7.

⁴⁸Ibid., p. 6.

⁴⁹Ibid.

Through such scoring the makers of the Short-Form have endeavoured to provide an instrument to be used for both survey and analytical purposes. It is suggested that educators, counsellors, psychologists, and employers can use the test in a wide variety of testing situations.⁵⁰

2. Assessment of validity and reliability. Most of the evidence of the validity of the Short-Form has attempted to prove adequate derived validity. Thus, the Short-Form has been scaled to the Stanford-Binet Intelligence Scale, 1963 Revision.⁵¹ Also, the Short-Form, Level 2, and the Stanford-Binet, 1963 Revision, when administered to the same subjects, showed a correlation of .87.⁵²

In addition, the Short-Form showed correlations of .94 and .88 with two group-administered mental ability tests.⁵³ With a measure of abstract intelligence as shown by achieved ability the Short-Form showed a correlation of .91.⁵⁴

Some evidence of direct validity has been given by

⁵⁰Ibid., p. 5.

⁵¹Ibid., p. 8.

⁵²Sullivan et al, Technical Report on the California Test of Mental Maturity Series, (Hereinafter referred to as Report), op. cit., p. 24.

⁵³Ibid., p. 126.

⁵⁴Ibid.



two groups of reviewers.⁵⁵ Both pointed out that the Short-Form had some observable defects. However, both felt that the test was useful, especially for lower grades or where pupils are still learning to read.

Both believed that the format of the test booklets was effective. Also, it appeared likely to the reviewers that a competent teacher could administer the tests.

Caution was advised in one review with regard to interpretation of factor scores.⁵⁶ However, for the present study only the total scores were used.

Finally, evidence which seemed to indicate good reliability has been presented. That reliability indicated that the test scores between administrations of Level 2 or the adjacent levels 2 and 3, with an interval of one year,⁵⁷ appeared to be highly consistent for individuals.

From the above summary of evidence it appears that the Short-Form, Level 2, has adequate validity and reliability as an I.Q. test. Thus, it has been used to assign pupils to I.Q. groups in the present study.

⁵⁵J.C. Stanley in O.K. Buros, Sixth Mental Measurements Yearbook, (Hereinafter referred to as Sixth), (Highland Park, N.J.: Gryphon Press, 1965), pp. 694-96; J.D. Linden and K.W. Linden, Tests on Trial (Boston: Houghton-Mifflin, 1968), pp. 22-24.

⁵⁶Linden and Linden, op. cit., p. 24.

⁵⁷.88 in both cases; as reported in Sullivan et al, Report, p. 19.

3. Use in the present study. Intelligence quotient had four levels as an independent variable in this study. These levels were determined from a system of descriptive categories suggested by the developers of the Short-Form.⁵⁸ The following levels were set up:

Level	Intelligence Quotient Range
High	116+
High Average	108 to 115
Average	93 to 107
Low	92-

It should be noted that the high level of the present study includes the high and very high categories from the interpretation guide.⁵⁹ Similarly the low level of the present study includes low average, low and very low from the interpretation guide.

The reason for combining extremely high or low categories was to provide several intelligence levels for comparison purposes while ensuring that each level was broad enough to yield a sample of close to thirty subjects. Thirty subjects in a sample was deemed sufficient to confidently

⁵⁸California Test Bureau Professional Staff, Guide to Interpretation of the California Test of Mental Maturity Series; 1963 Revision (Monterey, Calif.: California Test Bureau, 1964), p. 6; also, see Appendix C.

⁵⁹See Appendix C.

analyze data within the present study using the t-test technique.⁶⁰

As it turned out the low, average, high average, and high I.Q. groups consisted of 24, 51, 61, and 64 subjects, respectively. Thus only the smallest group included less than thirty subjects.⁶¹

Chronological Relationships Test. An attempt has been made in the following section to establish adequate validity and reliability for the Chronological Relationships Test which was used to assess pupil understanding of the chronological relationships embodied in the narrative and time line. The test, developed by this investigator, had to be specific to the materials used in the present study so that establishing direct validity by means of operational definitions appeared to be most appropriate.⁶² Reliability has been assessed in terms of Kuder-Richardson internal consistency methods.⁶³

1. Description. The Chronological Relationships Test consisted of 14 multiple-choice items. Each pupil's

⁶⁰Wiersma, op. cit., p. 75.

⁶¹The t-test was supplemented by the non-parametric median test for this group; see Appendix H.

⁶²Ebel, op. cit., pp. 383-85.

⁶³Ibid., pp. 318-30.

total score on the test has been used as data in this study. Also, the test had no time limit and all pupils were given sufficient time to complete it.

2. Validity. Under study in this investigation has been behavior which indicated the degree to which a pupil had understood or not understood the chronological relationships described in the narrative as well as shown on the time line.⁶⁴ The question to be answered in regard to validity is: To what extent do the Chronological Relationships Test items reflect the behavior under study?⁶⁵

To test such understanding a series of 14 multiple-choice test items was constructed by this investigator according to item construction criteria agreed upon by several writers on test construction.⁶⁶ The following test items measured the understandings through the behaviors described.

⁶⁴See Appendix A, Part 1.

⁶⁵Wiersma, op. cit., p. 191.

⁶⁶C.M. Lindvall, Measuring Pupil Achievement and Aptitude (New York: Harcourt, Brace and World, 1967), pp. 42-43; for similar suggestions see R.L. Ebel, Measuring Educational Achievement, pp. 151-70; R.H. Lindeman, Educational Measurement (Glenview, Ill.: Scott, Foresman and Co., 1967), pp. 76-83; N.E. Gronlund, Constructing Achievement Tests (Englewood Cliffs: Prentice-Hall, 1968), pp. 30-43.

UNDERSTANDINGS MEASURED BY THE
CHRONOLOGICAL RELATIONSHIPS
TEST ITEMS⁶⁷

Item Numbers	Understandings and Behaviors Demonstrating Understanding
1-8	Measured pupil understanding of the order of B.C. dates. The pupil had to apply the principle that earlier B.C. dates have larger numeric values in order to be certain that he was choosing the correct answer.
9-14	<p>Measured pupil understanding of how to compare at least four durations of time by:</p> <p>(a) computing each duration through finding the differences between the two numeric values of each duration's beginning and ending dates as well</p> <p>(b) comparing the resulting durations in order to choose the shortest or longest duration or durations.</p>

OR

⁶⁷See items in Appendix B.

Measured pupil understanding of how to compare at least four durations by:

(a) visually comparing the lengths of arrows showing duration on the time line, then choosing the shortest or longest durations as well as

(b) computing a duration by finding the difference between its beginning and ending dates.

The fourteen questions used were related to the chronological relationships expressed in a narrative and accompanying illustrative time line. "Ancient Civilizations,"⁶⁸ the title of the narrative and time line,⁶⁹ has been rated at grade six readability level on three readability scales: Dale-Chall, Flesch, and S.R.A.⁷⁰ In addition, as can be seen from inspection of the test, the vocabulary used in the test items should have presented little

⁶⁸Marius Livingstone, "Ancient Civilization" in N.B. Smith, Be A Better Reader, Foundations C, Teacher's Edition (Englewood Cliffs: Prentice-Hall, 1968), pp. 7-9.

⁶⁹See Appendix A, Part 1, for a reproduction of narrative and time line.

⁷⁰Smith, op. cit., p. T4; rated by the compilers of Smith, Be A Better Reader, Foundations C, op. cit.

difficulty at grade six level.⁷¹

Eleven specific dated events were included in the narrative and eight of these were shown on the illustrative time line. Questions on the Chronological Relationships Test pertained only to the order of and durations between the eight events so as to assure that the chronological relationships could be determined from the narrative or from the time line.

To summarize, this writer has attempted to establish the validity of the test of pupil understanding of chronological relationships through operational definitions of such understanding in terms of behaviors, described above, which appear to require such understanding. This investigator maintains examination of the items on the chronological relationships test is sufficient to establish validity of the items in terms of the above described behaviors. In addition, it appears likely that the greater the number of correct answers which each pupil chose the greater his understanding of the chronological relationships in the narrative or time line.

Furthermore, care has been taken to ensure that the vocabulary in the materials the test questions refer to, as

⁷¹Furthermore, reading comprehension was an independent variable in the study, thus providing a check on the effect of different reading comprehension levels on achievement on the test.

well as in the test itself, was approximately of grade six level. Finally, the questions were answerable by reference to either the narrative or the time line, thus allowing the experimental group to use the time line as an aid to understanding if they chose to.

3. Reliability. In addition to the reliability coefficient of the test, a report on the difficulty levels and discriminating powers of the individual items has been given below.

On the Chronological Relationships Test the item's difficulty indices ranged from .35 to .67⁷² with most items in the .40 to .65 range, indicating a good difficulty level⁷³ for nearly all items. Similarly, the items appeared to discriminate well⁷⁴ as the discrimination indices ranged from .55 to .89.⁷⁵

Finally, the reliability coefficient for the Chronological Relationships Test was .78,⁷⁶ which indicated a good reliability for a non-standardized achievement test.⁷⁷

⁷²See Appendix D.

⁷³Lindeman, op. cit., p. 91; Lindvall, op.cit., p.120.

⁷⁴Gronlund, op. cit., p. 87.

⁷⁵See Appendix D.

⁷⁶Using KR-20 formula; see Ebel, op. cit., pp. 318-20.

⁷⁷Gronlund, op. cit., p. 96.

4. Use in the present study. Each pupil's score, in terms of the number of items right on the Chronological Relationships Test, was obtained and used as a measure of his understanding of the chronological relationships incorporated into the narrative and on the illustrative time line.

DATA GATHERING PROCEDURES

After examining and assessing the validity and reliability of the test instruments used in the present study the procedures used in gathering data can now be specified.

Gathering Data on the Independent Variables

I.Q. and reading comprehension scores were taken from school district records. The I.Q. tests had been administered between December 15, 1969 and May 15, 1970. Since such scores have been shown to be relatively stable over up to a year's time⁷⁸ they were considered to be reliable for use in the present study, done in May and June, 1970. Reading comprehension tests had been administered in June 1970.

Chronological ages were computed from the cumulative records and sex was recorded by the pupils on their Chrono-

⁷⁸See p. 61.

logical Relationships Test answer sheets.

Applying Control and Experimental Treatments

In order to determine the amount of time required, a pilot project was done on May 5, 1970, in Edmonton, using a class of grade six pupils.⁷⁹ From that project it appeared that one hour was sufficient time for giving directions as well as administering and collecting the treatment materials and tests when allowing for even the slowest pupils to finish.

For the full study the treatment materials and the Chronological Relationships Test were administered on May 11 and May 12, 1970. In each class a full hour was available for administration.

Preliminaries. Upon entering each classroom the investigator gave the teacher the questionnaire on the teaching of chronology⁸⁰ and asked for the class register. The latter was used to choose pupils at random from the class to help pass out the materials.

Pupils were then requested to get a free-reading book each so that those finishing early would have something to

⁷⁹Selected by Edmonton Public School Board personnel.

⁸⁰For a copy of the questionnaire and a summary of the comments from it see Appendix E.

occupy their time. Also, any pupils who appeared to be sitting close enough to observe the details of other pupils' papers were requested to move their desks to more isolated positions.

At this point, five pupils were picked at random from the register and were given IBM answer sheets to hand out to the class. After the sheets were handed out the investigator checked with the class to see that each pupil had an HB pencil, as required for marking items to be machine-scored. On the answer sheets the pupils were asked to fill in their names, school and teacher, and whether they were male or female.

Directions for utilizing answer sheets. Since the pupils would have to read the questions in the test booklet then mark their answers on the answer sheet, directions for marking in answers were given, although when asked all pupils indicated familiarity with marking similar answer sheets. It was pointed out and demonstrated how the answers were to be marked in and pupils were cautioned to erase any changed answers completely so that the machine would not mark unintended answer spaces. Erasers were then made available for those pupils without their own.

To pass out the treatment materials five more pupils were selected randomly from those not selected previously. Before the materials were distributed the class was instructed to leave the treatment booklets face down on their desks

until given further instructions.

Random distribution of treatment materials. In order to ensure that the control (narrative only) and experimental (narrative and time line) treatment materials were assigned randomly, two steps were taken. First, the materials for each class have been previously stacked in mixed order, alternating between control and experimental treatment materials. Second, since the materials looked the same from the outside, the pupils unknowingly distributed the two treatments in a random manner.

Reading the treatment materials. After distribution of the treatment materials the pupils were asked to write their names on the fronts of the booklets. This was done so as to be able to later identify the treatment each pupil was on.

Since the Chronological Relationships Test was attached to the back of the treatment booklets, the pupils were asked to read only up to a stop sign at the bottom of page 3, which indicated the end of the treatment. In addition, they were asked to either free read or draw on the backs of the booklet pages they had read if they finished before everyone was done.

Instructions for doing the Chronological Relationships

Test. After all had completed reading the treatment materials, directions for marking the IBM answer sheets were reviewed. Also, in order to avoid undue anxiety, it was pointed out that their scores on the questions were not for school marks but for this investigator only.

In order to further allay anxiety the Chronological Relationships Test was always referred to as "the questions" not "the test." However, the pupils were requested to do the best they could and told that guessing at answers was acceptable if they were not sure of some answers.

Doing the Chronological Relationships Test. A final check was made to be sure all pupils had satisfactory pencils and erasers then the pupils were asked to turn to the fourteen questions starting on the page following the stop sign. They were told to turn back to any of the pages they had just read to help answer the questions.

When all had finished the Chronological Relationships Test the materials were collected by another group of pupil monitors. Before leaving each classroom the investigator picked up the questionnaire on time and chronology from the teacher.

Scoring the Chronological Relationships Test. All answer sheets were machine-scored on the IBM 1230 optical reader of the Division of Educational Research Services,

Faculty of Education, University of Alberta, Edmonton. Item analysis was carried out in order to determine a KR-20 reliability coefficient⁸¹ for the test as well as difficulty⁸² and discrimination⁸³ indices for each item.

NULL HYPOTHESES TESTED

In analyzing the data of the present study the following null hypotheses were tested.⁸⁴

1. At the .05 level⁸⁵ there is no significant difference between mean scores on the Chronological Relationships test in the experimental and control groups at the:

(a) high I.Q. level.

(b) high average I.Q. level.

(c) average I.Q. level.

(d) low I.Q. level.

2. At the .05 level there is no significant difference between mean scores on the Chronological Relationships

⁸¹See p. 68.

⁸²See p. 68.

⁸³See p. 68.

⁸⁴See hypotheses based on the research literature, Chapter 1, p. 7-8.

⁸⁵Significance level; see discussion in Wiersma, op. cit., pp. 70-74.

Test in the experimental and control groups at the:

- (a) high reading comprehension level.
- (b) average reading comprehension level.
- (c) low reading comprehension level.

3. At the .05 level there is no significant interaction between experimental or control group treatments and chronological age.

4. At the .05 level there is no significant difference between mean scores on the Chronological Relationships Test in the experimental and control groups for:

- (a) boys.
- (b) girls.

STATISTICAL PROCEDURES

Following the gathering of the data, the Chronological Relationships Test, I.Q., and reading comprehension scores as well as chronological age, sex, and treatment group data for each pupil were transferred to data cards for the IBM 360/67 computer.

Utilizing an ANOV1Ø computer program, t-tests were applied to test all except null hypothesis 3, as the latter was not a null hypothesis about the difference between two groups. A test for interaction using a MULRØ5 computer program was applied to test the latter hypothesis.

The T-test

Since the t-test is a parametric statistical test the following assumptions should be met:⁸⁶

Assumption of normality. This is apparently not the most important assumption behind the use of the t-test.⁸⁷ However, it is assumed in using the t-test that the samples with which one works have been drawn from populations that are normally distributed.

Homogeneity of variance. Here it is assumed that the variances within the groups are statistically the same.

Continuity and equal intervals as measures. The third assumption is that the measures to be analyzed are continuous measures with equal intervals. This assumption is behind the various arithmetic operations in the t-test.

In the case of the analyses carried out in the present study the above three assumptions appear to have been

⁸⁶Kerlinger, op. cit., pp. 258-60.

⁸⁷Ibid.; It is also claimed that "the empirical evidence suggests that even for quite small samples (5 or 10) reasonably large departures from normality will not seriously affect the estimation of probabilities for a two-tailed t-test" in Ferguson, op. cit., p. 145.

met. Evidence has been presented as follows:

First, the distribution of the total sample (all of the grade six pupils in the study) scores on the Chronological Relationships Test does not appear to depart dramatically from the normal distribution.⁸⁸ In addition, all samples are large enough, over ten, so that the assumption of normality is of minor importance.⁸⁹ As Kerlinger has pointed out:

Unless there is good evidence to believe that populations are rather seriously non-normal and that variances are heterogeneous, it is usually unwise to use a non-parametric statistical test in place of a parametric one. The reason for this is that the parametric tests are always more powerful⁹⁰ than non-parametric tests.⁹¹

In regard to assumption two, homogeneity of variance, it has been shown that the variances from group to group are "within the bounds of random variation."⁹² Appendix G compares the variance within each experimental group with each corresponding control group and reports the probability that the differences are random variations, thus indicating

⁸⁸See histogram in Appendix F.

⁸⁹Ferguson, op. cit., p. 145.

⁹⁰The power of a statistical test is the probability that the null hypothesis will be rejected when it is actually false.

⁹¹Kerlinger, op. cit., p. 259.

⁹²Ibid., p. 258.

homogeneous variances.

Apparently, assumption three has been met also. The Chronological Relationships Test scores have been determined on a fifteen point continuous measure with fourteen equal intervals.⁹³

From the above evidence it appears likely that the t-test has been suitably used in the present study. In the next section, why and how the t-test has been used in this study has been examined.

First, for all null hypotheses except hypothesis three the means of a control and an experimental group have been compared. The t-test is essentially a test which determines what the probability is that the difference between a pair of means is simply due to random variation rather than any experimental effect.

Within the present study, since experimental and control treatment effects were hypothesized to be statistically equal, two-tailed t-tests⁹⁴ were used on all null hypotheses tested with the t-test except null hypothesis 1.a. In that hypothesis it had been predicted earlier⁹⁵ that the

⁹³From a fourteen item test with any score from zero through fourteen possible.

⁹⁴Wiersma, op. cit., pp. 83-85; Ferguson, op. cit., pp. 135-36; G.H. Weinberg and J.A. Schumaker, Statistics: An Intuitive Approach, pp. 162-65.

⁹⁵See Chapter 2, p. 31.

experimental group would achieve better, thus making a one-tailed t-test appropriate since the latter test may be used only when it has been hypothesized that a specified group will achieve better than the other group.⁹⁶

Test for Interaction

In applying a test for interaction⁹⁷ the following question has been answered: Is there a constant difference on criterion test⁹⁸ performance between the two treatment conditions⁹⁹ across the observed chronological ages?

Program MULRØ5 has been used to test for any interaction affect.¹⁰⁰ The likelihood of interaction between treatment and chronological age has been computed in terms of a probability figure indicating whether or not to reject a null hypothesis of no significant difference¹⁰¹ between the treatment groups was computed. For purposes of the

⁹⁶Wiersma, op. cit., pp. 83-85.

⁹⁷discussed in F.J. Kelly, D.L. Beggs, and K.A. McNeil, Research Design in the Behavioral Sciences: Multiple Regression Approach (Carbondale, Ill.: Southern Illinois University Press, 1969), pp. 139-44.

⁹⁸Chronological Relationships Test.

⁹⁹experimental and control.

¹⁰⁰Division of Educational Research Services, Faculty of Education, University of Alberta, Edmonton.

¹⁰¹in the Chronological Relationships Test scores.

present study, as indicated earlier for the t-test,¹⁰² the significance level for rejection of the null hypothesis of no significant interaction is .05.¹⁰³

In addition to the above probability figures the MULRØ5 program calculated constants for use in a prediction equation. That equation was used in making a graph¹⁰⁴ showing the relationship of age to criterion score in each treatment group. These lines of best fit for the data indicated whether the criterion scores in the treatment groups tended to increase or decrease as chronological age increased.

SUMMARY

Within the present chapter the methodology of the present study has been elaborated as follows:

First, the posttest-only control group design has been explained and justified.

Second, the dependent and independent variables have been reviewed.

Fourth, the tests used to gather data on the three

¹⁰²See p. 74.

¹⁰³See a discussion of determination of significance level in Wiersma, op. cit., pp. 73-74.

¹⁰⁴See a discussion of plotting the lines of best fit in Kelly et al, op. cit., pp. 150-55.

variables of I.Q., reading comprehension, and understanding of chronological relationships have been described, assessed and related to the present study.

Fifth, the specific data-gathering procedures used have been specified.

Sixth, a number of null hypotheses which depended on the data for acceptance or rejection have been elaborated.

Finally, the statistical procedures used to test the hypotheses have been explained.

Chapter 4

RESULTS

Within this chapter each null hypothesis has been restated and the results of the analysis of data pertaining to each has been examined. On the basis of the probability associated with each null hypothesis it has been either rejected or not rejected. The null hypotheses and data pertaining to them are grouped according to the four independent variables of I.Q., reading comprehension, age, and sex.

NULL HYPOTHESES RELATING TO INTELLIGENCE QUOTIENT

First, the null hypothesis relating to intelligence quotient levels have been examined. As mentioned earlier,¹⁰⁵ two tailed tests were used in testing all hypotheses except hypothesis 1.a., which will be dealt with first.

¹⁰⁵in statistical procedures, Chapter 3.

Null hypothesis 1.a. stated: At the .05 level there is no significant difference between mean scores on the Chronological Relationships Test for the experimental and control groups at the high I.Q. level.

Within the high I.Q. sample, as shown in Appendix H, the experimental group had a mean score of 9.94 while the control group had a mean score of 7.29. Upon applying the t-test it was found that the difference between the means had a probability of less than the .05 level. Accordingly, null hypothesis 1.a. was rejected.

Null hypothesis 1.b. stated: At the .05 level there is no significant difference between mean scores on the Chronological Relationships Test for the experimental and control groups at the high average I.Q. level.

As displayed in Appendix H, within the high average I.Q. sample the subjects in the experimental and control groups had mean scores of 8.59 and 6.35, respectively. Application of a t-test indicated that the difference between the means had a probability of less than .05. Consequently, null hypothesis 1.b. was rejected.

Null hypothesis 1.c. stated: At the .05 level there is no significant difference between mean scores on the Chronological Relationships Test for the experimental and control groups at the average I.Q. level.

In the average I.Q. sample, the mean scores were 6.08, experimental group, and 4.92, control group, as indi-

cated in Appendix H. In this case a t-test indicated that a probability of more than .05 was associated with the difference between the means. As a result, null hypothesis 1.c. could not be rejected.

Null hypothesis 1.d. stated: At the .05 level there is no significant difference between mean scores on the Chronological Relationships Test for the experimental and control groups at the low I.Q. level.

Mean scores for the experimental and control groups were 4.70 and 4.66, respectively, in the low I.Q. sample. A probability of more than .05 was associated with the difference between the means, according to a t-test. Hence, null hypothesis 1.d. was not rejected.

NULL HYPOTHESES RELATING TO READING COMPREHENSION LEVELS

Having dealt with the null hypotheses relating to intelligence levels, those relating to reading comprehension have been evaluated next. Two-tailed tests of significance were applied to the three reading hypotheses, as pointed out earlier.¹⁰⁶

Null hypothesis 2.a. stated: At the .05 level there is no significant difference between mean scores on the Chrono-

¹⁰⁶in statistical procedures, Chapter.3.

logical Relationships Test for the experimental and control groups at the high reading comprehension level.

In the high reading comprehension sample, as indicated by Appendix I, the experimental group had a mean score of 9.98 while the control group had a mean score of 7.53. A t-test indicated that the difference between the means had a probability of less than .05. Thus, null hypothesis 2.a. was rejected.

Null hypothesis 2.b. stated: At the .05 level there is no significant difference between mean scores on the Chronological Relationships Test for the experimental and groups at the average reading comprehension level.

The experimental and control groups had mean scores of 7.34 and 5.47, respectively, as shown in Appendix I. A probability of less than .05 was determined for the difference between the means. Therefore, null hypothesis 2.b. was rejected.

Null hypothesis 2.c. stated: At the .05 level there is no significant difference between mean scores on the Chronological Relationships Test for the experimental and control groups at the low reading comprehension level.

Mean scores for the experimental and control groups were 5.43 and 3.77, respectively. Using a t-test, a probability of less than .05 was computed for the difference between the means. Accordingly, null hypothesis 2.c. was rejected.

NULL HYPOTHESIS RELATING TO AGE

A further null hypothesis relating to chronological age has been tested. This hypothesis involved a test for interaction.

Null hypothesis 3 stated: At the .05 level there is no significant interaction between experimental or control group treatments and chronological age.

The probability of interaction was calculated to be .12 which indicated interaction was not significant at the .05 level. Thus, null hypothesis 3, which indicated no significant interaction, was not rejected.

Some other figures of interest were computed also. First, the mean scores of control and experimental groups were 8.04 and 6.04, in that order, and the difference between the means was significant at .001, using an F-test.¹⁰⁷ Second, as shown in Appendix J, for both experimental and control treatments Chronological Relationships Test scores tended to decrease as chronological age increased.

In order to see if increased mental age might be a better predictor of increased scores on the Chronological Relationships Test, the following a posteriori null hypothesis, which indicated no significant interaction, was not rejected.

¹⁰⁷See Weinberg and Schumaker, op. cit., pp. 309-15.

Further, as shown in Appendix K, Chronological Relationships Test scores for the experimental group tended to increase slightly as mental age increased. However, there appeared to be no increase in scores for the control group as mental age increased. More generally, it appeared that as mental age increased Chronological Relationships Test scores did not increase significantly for either the experimental or the control group.

NULL HYPOTHESES RELATING TO SEX

Two final null hypotheses, relating to sex, were tested within the present study. Once again two-tailed tests of significance were applied.

Null hypothesis 4.a. stated: At the .05 level there is no significant difference between mean scores on the Chronological Relationships Test in the experimental and control groups for boys.

Through reference to Appendix L, it can be seen that 8.66 and 6.36 were the mean scores for boys in the experimental and control groups, respectively. Using a t-test, it was shown that the difference between means had a probability of less than .05. As a result, null hypothesis 4.a. was rejected.

Null hypothesis 4.b. stated: At the .05 level there is no significant difference between mean scores on the

Chronological Relationships Test in the experimental and control groups for girls.

For girls, the mean scores in the experimental and control groups were 7.22 and 5.65, in that order. A probability of less than .05 was determined for the difference between the means, using a t-test. It followed that null hypothesis 4.b. had to be rejected.

An additional analysis was carried out on the data for boys and for girls. It appeared to be of some interest whether there was a significant difference between boys and girls on each separate treatment.

Accordingly, a t-test was performed on the difference between the boys' and girls' means within the experimental group. The probability computed was .40 which indicated that at the .05 level there was no significant difference between the means.

Similarly, within the control group, a probability of .28 was associated with the difference between the boys' and girls' means. Once again, at the .05 level, no significant difference between the means was shown.

SUMMARY

The results of the tests were as follows:

First, there were significant differences between the mean Chronological Relationships Test scores of the experimental and control groups, favoring the experimental group,

at the high and high average I.Q. levels. However, at the average and low I.Q. levels the mean scores of the control and experimental groups were not significantly different.

Second, at low, average, and high reading comprehension levels there were significant differences between the Chronological Relationships Test mean scores of the experimental and control groups, with the former group excelling.

Third, it appeared that there was no significant interaction between chronological or mental age and treatment. Also, increased chronological or mental age did not appear to be good predictors of increased scores on the Chronological Relationships Test.

Fourth, there was a significant difference between Chronological Relationships Test means on the experimental and control groups for boys as well as for girls. As an additional point of interest, it was calculated that there was no significant difference between the means of boys and girls within either the experimental or the control group.

Chapter 5

CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

REVIEW OF THE PROBLEM AND PROCEDURES

Within this study an attempt has been made to answer the following question: Is the effect of an illustrative time line with a historical narrative superior to the historical narrative alone in terms of pupil understanding of the chronological relationships in the narrative?

In order to guide the investigation of the above question, four hypotheses about the affect of an illustrative time line, based on relevant research literature, were stated. Those hypotheses were as follows:

In terms of pupil achievement on a chronological relationships test the effect of a historical narrative with a time line:

1. is superior to the effect of the narrative alone at a high intelligence level.
2. is the same as the effect of the narrative alone at low, average, or high reading comprehension levels.
3. increases equally with the effect of the nar-

rative alone as chronological age increases.

4. is the same as the effect of the narrative alone for boys or girls.

To gather data indicating support or non-support of the hypotheses, an experiment was set up at the grade six level, using 214 subjects assigned randomly to two treatment groups. Treatment materials were selected and a fourteen item chronological relationships test was developed to assess pupil understanding of the chronological relationships incorporated into the materials.

For the experimental and control groups, respectively, the treatment materials consisted of a historical narrative with an illustrative time line and the historical narrative by itself. Eight dated events were incorporated into the narrative as well as shown on the time line.

Data consisting of scores on the chronological relationships test, I.Q. scores, reading comprehension scores, chronological and mental ages, and sex was collected. A number of null hypotheses derived from the four directive hypotheses were tested through analysis of the data. The conclusions of the present chapter are based on the results of the hypothesis-testing reported in chapter four.

CONCLUSIONS

On the basis of the findings reported in chapter four

the following conclusions, to be discussed in the next section, have been drawn. The first is the general conclusion of the study and the others more specific conclusions.

1. The effect of the illustrative time line with the historical narrative was significantly superior to the historical narrative alone in terms of pupil understanding of the chronological relationships in the narrative.

2. The effect of the historical narrative with the time line was significantly superior to the effect of the narrative alone at high as well as high average intelligence levels. However, at average and low intelligence levels the effect did not appear to be significantly different for the experimental and control groups.

3. At high, average, and low reading comprehension levels the effect of the narrative with the time line significantly surpassed the effect of the narrative alone.

4. Increased chronological age was not a good predictor of increased understanding of the chronological relationships in the narrative for either those with or those without the illustrative time line. Understanding appeared to decrease slightly at higher chronological age levels. However, across all chronological age levels the effect of the narrative with the time line was significantly superior to the effect of the narrative alone.

Increased mental age was not a good predictor of significant improvement in understanding of chronological

relationships either, although understanding tended to increase slightly as mental age increased.

5. For boys or girls the effect of the narrative with the time line was significantly superior to the effect of the narrative alone. Furthermore, the narrative with time line effect was not significantly different for boys and girls. The same was true for the narrative by itself. Apparently, the sex of the subjects did not significantly influence effective utilization of the time line.

DISCUSSION

The conclusions of this study will now be discussed in relation to the findings of the earlier studies reviewed in chapter two. Since the general conclusion depends upon the specific conclusions only the latter are dealt with below.

First, the conclusions of the present study appear to be supported by the earlier finding by Davis that at the high intelligence level pupils appeared to benefit from an illustrative time line.¹⁰⁸ However, Davis found that at his I.Q. level II, approximately equivalent to the high average level of the present study,¹⁰⁹ the pupils did not benefit signif-

¹⁰⁸Davis et al, op. cit., pp. 22-25.

¹⁰⁹108-115 I.Q. range versus Davis' 110-119 I.Q. range.

icantly from the time line.

The latter contradiction between Davis' and the present study may result from differing procedures. As mentioned earlier, Davis' study appeared to require more dependence on memory and less on intelligence on the part of the subjects, when compared to the present study.¹¹⁰ Thus, higher intelligence pupils may have had more opportunity to exercise their intelligence when utilizing the time line in the present study.

Another contradiction arises from earlier studies which had found that it is not possible to infer that illustrations affect reading comprehension.¹¹¹ The conclusion of the present study that an illustrative time line can aid comprehension of chronological relationships described in a written narrative seems to contradict the earlier findings.

An explanation of the discrepancy between the findings may be attributable to the types of illustrative materials used. The cross-sectional drawings, story-book pictures, and graphs¹¹² used in the earlier studies illustrated relatively concrete objects such as valves, children, or money as compared to the time line in the present study which illustrated

¹¹⁰See Chapter 2, pp. 14-15.

¹¹¹Burdick, op. cit., p. 37; Vernon, op. cit., p. 186; Miller, op. cit., pp. 181-82.

¹¹²Ibid.

an abstract concept, time.

An abstract concept such as time may be more difficult for children to visualize than concrete objects. Illustrations of concrete objects may not have produced any significant increase in the children's comprehension of the written materials describing the objects because the children may have been capable of visualizing such objects with sufficient accuracy, in the earlier studies. However, in the present study, those using the time line may have derived benefit from being able to see an illustration of the relatively abstract concept, time, which those without the time line may have visualized less successfully.

A further contradiction between this study and earlier studies was indicated by the present conclusion that increased chronological age is associated with a slightly decreased understanding of chronological relationships, regardless of whether the time line is present. A number of earlier studies had concluded that increased chronological age was a predictor of increased ability to deal with time and chronology.¹¹³ There seems to be at least two possible reasons why chronological age was not such a predictor within the present study.

First, the present study encompassed only the chronological age range within a sixth grade sample. The earlier

¹¹³Oakden and Sturt, op. cit., pp. 334-36; Kelty, op. cit., pp. 615-18; Harrison, op. cit., pp. 513-14; Bradley, op. cit., pp. 77-78; Callahan in Chase, "American History," op. cit., pp. 337-38; Farrel, op. cit., pp. 592-94; Gill, op. cit., pp. 455-56; Newman, op. cit., pp. 96-97.

studies¹¹⁴ all included several grades, consequently providing a broader chronological age range. Such a wider range may have resulted in a more accurate picture of the relationship between chronological age and changing achievement. The present study may present an inaccurate picture because the range of chronological age is not great enough even though the sample size is larger than those in some of the contradictory studies.¹¹⁵

A second reason why the older pupils appeared to be less able to understand chronological relationships in the present study may reside in the fact that some of the grade six pupils, particularly the older ones, were held over once or twice in the earlier grades.¹¹⁶ Hence, across the chronological age levels from youngest to oldest, there likely tended to be an increasing number of hold overs compared to non-hold overs.

Any pupils who were held over would generally be those of lower intelligence and weaker reading comprehension skills¹¹⁷ so that they would probably be less able to under-

¹¹⁴Ibid.

¹¹⁵Harrison, op. cit., pp. 507-08; Farrel, op. cit., p. 588; Newman, op. cit., p. 93.

¹¹⁶The school system has not used a continuous progress plan.

¹¹⁷Intelligence and reading comprehension were two of the criteria used in determining who to hold back.

stand concepts of time and chronology. Thus, at higher chronological age levels accomplishment on the chronological relationships test tended to decrease slightly.

Since increased chronological age was not a good predictor of increased understanding of chronological relationships, mental age was also investigated as such a predictor. However, increased mental age did not appear to be a predictor of significant improvement in understanding of chronological relationships either, although understanding increased slightly as mental age increased. As with chronological age, it may be that the range of mental age within the present study was not broad enough to provide an accurate picture of the relationship between mental age and changing achievement.

A more pronounced relationship might have become evident if the range of mental age had been over several grades as in Harrison's study.¹¹⁸ She had found a high correlation between understanding of concepts of time and mental age, for pupils in kindergaten through grade three.

Finally, the conclusion that the sex of the subjects did not significantly influence ability to effectively use the time line is supported by the findings of some earlier

¹¹⁸Harrison, op. cit., pp. 513-14.

studies¹¹⁹. In addition, the findings by Callahan¹²⁰ and Dobbs¹²¹ that boys and girls excell at dealing with differing aspects of time and chronology did not appear to significantly affect the findings of the present study.

LIMITATIONS

Certain limitations of the present study must be considered in any interpretation of the conclusions just discussed. At least two limitations seem to be particularly pertinent.

First, the study has been based on a sample consisting of the grade six pupils from one small city in Alberta. Thus, anyone wishing to generalize to other groups of grade six pupils will have to judge how closely his sample and situation approximates that of this study.

Second, only one type of time line and one historical narrative have been used in this study. Perhaps the use of other types of time lines or other narratives would yield different results.

¹¹⁹Friedman, op. cit., pp. 337-42; Farrel, op. cit., pp. 587-94; Newman, op. cit., p. 96.

¹²⁰Callahan in Chase, "American History," op. cit., pp. 337-38.

¹²¹Dobbs, op. cit., p. 85.

IMPLICATIONS

Perhaps the major implication of the present study is the support given to the inclusion of illustrative time lines in historical text materials. It has been shown in this study that at least one type of illustrative time line appears to be useful at grade six level, thus supporting the claims¹²² that such time lines can be of value to elementary school pupils.

Some earlier studies¹²³ which concluded that illustrations are of no value are not supported by this study. It appears that at least one specific type of illustration, the time line as used in this study, aids comprehension of written descriptions of related chronological relationships.

Finally, the finding that pupils can put an illustrative time line to use suggests that they might benefit from making their own time lines as well. If pupils can derive benefit from a ready-made illustrative time line they may derive at least some benefit from accurate time lines constructed by themselves. Constructing unique time lines may be a good way to develop understanding of the principles underlying time lines. On the other hand, investigation may

¹²²See Chapter 1, p.2.

¹²³See Chapter 2, p.16.

show that pupils derive no benefit from constructing their own time lines.

RECOMMENDATIONS FOR RESEARCH

Several recommendations for further formal or informal research have been suggested below. These recommendations are not necessarily in order of importance.

First, an investigation of the effect of short and long-term training, over several weeks, months, or years, in how to use time lines could be carried out. Such a study might investigate whether average or low I.Q. pupils particularly can be taught to utilize a time line efficiently and with significant effect.

Second, studies of the effectiveness of various types of time lines might be carried out. For example, instead of using arrows to indicate durations as in the present study, durations might be shown in other ways or left for pupils to determine from the positions of events on the time line.

Third, further studies of the effectiveness of time lines could be carried out in lower elementary grades. The present study has dealt only with a grade six sample.

Fourth, studies of the effectiveness of time lines when combined with teacher or pupil oral explanations and discussions might be done. In the present study all information was presented in written form or on the time line.

Fifth, investigations of the effectiveness of time lines which add to rather than repeat chronological information presented in the accompanying written material could be carried out. For example, instead of presenting the same precise dates of events on both the time line and in the narrative, as in the present study, the time line might be used to convey precise dates while the narrative might only use imprecise expressions such as "several centuries later," "many years before" and so on.

Finally, the relationship between such factors as numerical reasoning and spatial reasoning ability and ability to benefit from a time line might be investigated. Numerical reasoning ability probably affects the ability to comprehend the order of dates and compute durations, especially when a pupil does not have access to a time line. In addition, spatial ability probably affects a pupil's ability to compare durations by comparing segments on a time line.

BIBLIOGRAPHY

- Ames, L.B. "The Development of the Sense of Time in the Young Child," Pedagogical Seminary and Journal of Genetic Psychology, LXVIII (April, 1946), 97-125.
- Arnsdorf, V.E. "An Investigation of the Teaching of Chronology in the Sixth Grade," Journal of Experimental Education, XXIX (March, 1961), 307-13.
- Blishen, B.R. "The Construction and Use of an Occupational Scale,": Canadian Journal of Economics and Political Science, XXIV (November, 1958), 519-31.
- Bradley, N.C. "The Growth of the Knowledge of Time in Children of School Age," British Journal of Psychology, XXXVIII (1947), 67-78.
- Burdick, J.G. "A Study of Cross-Section Drawings as Technical Illustrations in High School Textbooks.: Unpublished Ph. D. dissertation, Syracuse University, 1959.
- Buros, O.K. Fifth Mental Measurements Yearbook. Highland Park, N.J.: Gryphon Press, 1959.
- Buros, O.K. Sixth Mental Measurements Yearbook. Highland Park, N.J.: Gryphon Press, 1965.
- California Test Bureau Professional Staff. Guide to Interpretation of the California Test of Mental Maturity Series. Monterey, Calif.: California Test Bureau, 1963.
- Campbell, D.T., and J.C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNalley, 1963.
- Chase, W.L. A Guide for the Elementary Social Studies Teacher. Boston: Allyn and Bacon, 1968.
- Davis, O.L., L.C. Hicks, and N.D. Bowers. "The Usefulness of Time Lines in Learning Chronological Relationships in Text Materials," Journal of Experimental Education, XXXIV (Spring, 1966), 22-25.
- Davis, O.L. "Usefulness of a Time Line - With Historical Text," Educational Leadership, XXV (May, 1968), 750-53.
- Dobbs, E.L. "A Study of the Sequential Development of Time Sense and Chronology in the Elementary School," Unpublished Ed. D. thesis, University of Kansas, 1963.

- Ebel, R.L. Measuring Educational Achievement. Englewood Cliffs: Prentice-Hall, 1965.
- Education Index. New York: H.W. Wilson.
- Farrell, M. "Understanding of Time Relations of Five-, Six-, and Seven-year old Children of High I.Q.," Journal of Educational Research, XLVI (April, 1953), 587-94.
- Ferguson, G.A. Statistical Analysis in Psychology and Education. New York: McGraw-Hill, 1959.
- Friedman, K.C. "Time Concepts of Elementary-School Children," Elementary School Journal, XXXIV (February, 1944), 337-42.
- Friedman, K.C. "Time Concepts of Junior and Senior High School Pupils and of Adults," School Review, LII (April, 1944), 233-38.
- Friedman, K.C., and V.A. Marti. "A Time Comprehension Test," Journal of Educational Research, XXXIX (Sept., 1945), 62-68.
- Gill, C.G. "Interpretations of Indefinite Expressions of Time," Social Education, XXVI (December, 1962), 454-56.
- Gronlund, N.E. Constructing Achievement Tests. Englewood Cliffs: Prentice-Hall, 1968.
- Harrison, M.L. "The Nature and Development of Concepts of Time Among Young Children," Elementary School Journal, XXXIV (March, 1934), 507-14.
- Jarolimek, J. Social Studies in Elementary Education. 3rd ed. Toronto: Collier-Macmillan, 1967.
- Kelly, T.J., D.L. Beggs, and K.A. McNeil. Research Design in the Behavioral Sciences: Multiple Regression Approach. Carbondale, Ill.: Southern Illinois University Press, 1969.
- Kelty, M.G. "Time Expressions Comprehended by Children in the Elementary School," Elementary School Journal, XXV (April, 1925), 607-18.
- Kerlinger, F.N. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston, 1964.
- King, E.M., ed. Canadian Tests of Basic Skills, Form 2. Toronto: Thomas Nelson and Sons, 1967.

- Lindeman, R.H. Educational Measurement. Glenview, Ill.: Scott, Foresman, 1967.
- Linden, J.D., and K.W. Linden. Tests on Trial. Boston: Houghton-Mifflin, 1968.
- Lindvall, C.M. Measuring Pupil Achievement and Aptitude. New York: Harcourt, Brace and World, 1967.
- Lundberg, G.A., C.C. Schrag, and O.N. Larsen, Sociology. rev. ed. New York: Harper and Row, 1963.
- McAuley, J.D. "What Understanding Do Second Grade Children Have of Time Relationships," Journal of Educational Research, LIV (1961), 312-14.
- Miller, W.A. "Reading With and Without Pictures," Elementary School Journal, XXXVIII (May, 1938), 678-82.
- National Council for the Social Studies. Improving the Teaching of World History. Twentieth Yearbook. Washington: N.C.S.S., 1949.
- National Council for the Social Studies. Interpreting and Teaching American History. Thirty-first Yearbook. Washington: N.C.S.S., 1961.
- National Council for the Social Studies. Skill Development in the Social Studies. Thirty-third Yearbook. Washington: N.C.S.S., 1963.
- Newman, W.O. "Childrens' Understanding of Time Duration." Unpublished Master's thesis, University of Alberta, Edmonton, Canada, 1967.
- Oakden, E.C., and M. Sturt. "The Development of the Knowledge of Time in Children," British Journal of Psychology, XII (April, 1922), 334-36.
- Pistor, F. "Measuring the Time Concepts of Children," Journal of Educational Research, XXXIII (December, 1939), 293-300.
- Pistor, F. "How Time Concepts Are Acquired By Children," Educational Method, XX (November, 1940), 107-12.
- Preston, R.C. Teaching Social Studies in Elementary School. 3rd ed. Toronto: Holt, Rinehart and Winston, 1968.
- Russel, D.H. Children's Thinking. Toronto: Ginn and Co., 1956.

- Smith, N.B. Be A Better Reader, Foundations C. Teacher's Edition. Englewood Cliffs: Prentice-Hall, 1968.
- Sturt, M. The Psychology of Time. London: Routeledge and Kegan Paul, 1922.
- Sullivan, E.T., W.W. Clark and E.W. Tiegs. California Short-Form Test of Mental Maturity. Monterey, Calif.: California Test Bureau, 1963.
- Sullivan, E.T., W.W. Clark, and E.W. Tiegs, California Test of Mental Maturity. Monterey, Calif.: California Test Bureau, 1963.
- Sullivan, E.T., W.W. Clark, and E.W. Tiegs. Technical Report on the California Test of Mental Maturity Series. Monterey, Calif.: California Test Bureau, 1963.
- U.S. Office of Education. Educational Resources Information Center. Washington: U.S. Office of Education.
- Vernon, M.D. "The Value of Pictorial Illustration," British Journal of Educational Psychology, XXIII (1953), 180-87.
- Webster's Seventh New Collegiate Dictionary. Toronto: Thomas Allen and Sons, 1969.
- Weinberg, G.H., and J.A. Schumaker. Statistics: An Intuitive Approach. Belmont, Calif.: Wadsworth, 1969.
- Wesley, E.B., and M.A. Adams. Teaching Social Studies in Elementary Schools. rev. ed. Boston: D.C. Heath, 1952.
- Wiersma, W. Research Methods in Education. New York: J.B. Lippincott, 1969.

APPENDICES

APPENDIX A
TREATMENT MATERIALS

APPENDIX A, PART 1
EXPERIMENTAL TREATMENT

NAME _____

SOME EARLY CIVILIZATIONS

No one knows just when men first developed civilizations, but scientists now think that it was near the end of the prehistoric period. Prehistoric means before history. The prehistoric period extends from the beginning of time to about 3000 B.C.

About 3000 B.C. men invented a way of writing down their language. They left written records of their civilizations. With the development of a written language, the prehistoric period ends, and history begins.

Early prehistoric people did well just staying alive. They did not develop a civilization. However, we know that from about 3000 B.C. to the birth of Christ several groups of people did develop important civilizations. They left written records of their civilizations.

The Egyptians, Babylonians, Assyrians, and the Greeks were among the ancient peoples who developed such civilizations. In their civilizations they had a language, and organized government, and religion. Arts and skills of many kinds were also part of these civilizations.

THE EGYPTIANS

The Nile River formed a valley in northern Africa. The people who lived in this valley were called Egyptians (i jip'shənz). Soil in this valley was fertile and life was easier than in many other places in the world.

The Egyptians were among the first to organize a formal government, to develop a written language, and to use metals. They also gave the world some new information about geometry, chemistry, and medicine.

The Egyptian civilization extended from about 3000 B.C. to 671 B.C. In 671 B.C. Egypt was defeated by the Assyrians.

THE BABYLONIANS

The Babylonian (bab' ə lō'ne ən) civilization developed between the Tigris (tī'gris) and Euphrates (ū frā'tēz) rivers by about 2700 B.C. This great civilization was powerful for about for almost a thousand years. By 1900 B.C. the Babylonians had conquered most of the peoples in the Tigris-Euphrates Valley. About 910 B.C. Babylonia was conquered by the Assyrians.

The Babylonians were the first to make use of the arch in brick buildings. They irrigated their lands and improved the ways of making ditches and canals. They studied the stars and made a calender. Their writing system was very unusual. It was called cuneiform (Kū nē'fōrm), which means wedge-shaped.

* narrative without time line

THE ASSYRIANS

The Assyrians lived north of Babylonia, on the Tigris River. They developed a civilization by about 2100 B.C. At one time they were a peaceful people and stayed in their own land. Later they became fierce and warlike. With huge armies they swept down on neighboring cities. Armed with iron spears and swords, they struck terror into the hearts of other peoples.

In 910 B.C. the Assyrians conquered Babylonia. By 625 B.C. they had built a huge empire of captured peoples. This empire included not only the people in the Tigris-Euphrates Valley, but the Egyptians as well.

But Assyria became weak from constant warfare. Rather suddenly this great empire came to an end. Between 612 and 606 B.C. Assyria was destroyed by invaders from the south and the east.

While the Assyrians were in power they developed trade to distant lands. They built great roads and cities. One of their kings brought together a famous library of clay tablets. They tell us much about the early civilizations.

THE GREEKS

The Greeks lived in a smaller area of land than the other peoples we have discussed. Furthermore, this land was mountainous, rocky, and certainly not fertile. For this reason, many Greeks moved to nearby islands and to fertile coastlands in distant places. Many of them became traders.

Most of the ancient Greeks lived in city-states. These were cities which were like small, independent countries in themselves. The Greek people were loyal to their own city-state. There were hundreds of these city-states in ancient Greece.

The ancient Greek civilization began about 1200 B.C. It ended as such about 338 B.C. By 338 B.C. Phillip of Macedonia and his famous son, Alexander the Great, had taken over the Greek city-states. These states had been weakened by fighting and jealousy among themselves. One by one, the city-states lost their independence and became part of Alexander's kingdom.

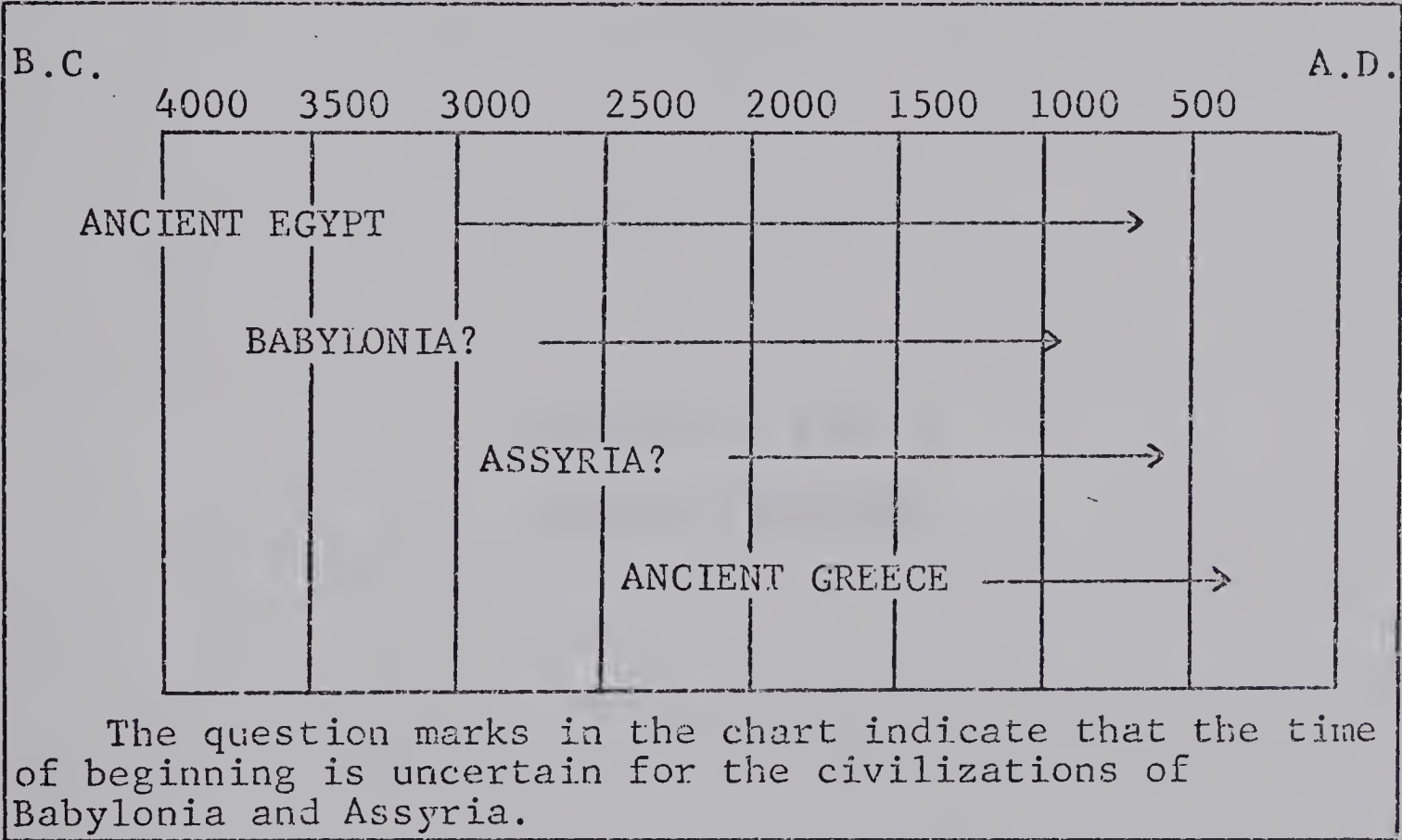
In spite of their small land area and rocky soil, the ancient Greeks developed a civilization which is most important to us. They were the first people to practice democracy (di mok'rə sē). They were the first to have the Olympic (ō lim'pik) Games. Many of our ideas in the fields of government, literature, and art have come from the ancient Greeks.

** The phonetic spellings remain unchanged from Livingstone's "Ancient Civilizations" in N.B. Smith, Be a Better Reader, Foundations C, Teacher's Edition (Englewood Cliffs: Prentice-Hall, 1968), pp. 7-9.

TIME LINE

The time covered in the article which you have just read is from 4000 B.C. to the beginning of A.D. Dates extending from B.C. to A.D. are read backward. Thus, 1000 B.C. means 1,000 years before the birth of Christ, and 2000 B.C. means 2,000 years before the birth of Christ.

A.D. dates are read forward. Thus, 500 A.D. means 500 years after the birth of Christ, and 1000 A.D. means 1,000 years after the birth of Christ. The article you read discussed ancient civilization only. So all of your work in the following questions will be with B.C. time.



APPENDIX A, PART 2
CONTROL TREATMENT

NAME _____

SOME EARLY CIVILIZATIONS

No one knows just when men first developed civilizations, but scientists now think that it was near the end of the prehistoric period. Prehistoric means before history. The prehistoric period extends from the beginning of time to about 3000 B.C.

About 3000 B.C. men invented a way of writing down their language. They left written records of their civilizations. With the development of a written language, the prehistoric period ends, and history begins.

Early prehistoric people did well just staying alive. They did not develop a civilization. However, we know that from about 3000 B.C. to the birth of Christ several groups of people did develop important civilizations. They left written records of their civilizations.

The Egyptians, Babylonians, Assyrians, and the Greeks were among the ancient peoples who developed such civilizations. In their civilizations they had a language, and organized government, and religion. Arts and skills of many kinds were also part of these civilizations.

THE EGYPTIANS

The Nile River formed a valley in northern Africa. The people who lived in this valley were called Egyptians (i jip'shənz). Soil in this valley was fertile and life was easier than in many other places in the world.

The Egyptians were among the first to organize a formal government, to develop a written language, and to use metals. They also gave the world some new information about geometry, chemistry, and medicine.

The Egyptian civilization extended from about 3000 B.C. to 671 B.C. In 671 B.C. Egypt was defeated by the Assyrians.

THE BABYLONIANS

The Babylonian (bab' ə lō nē' ən) civilization developed between the Tigris (tī'gris) and Euphrates (ū frā'tēz) rivers by about 2700 B.C. This great civilization was powerful for about almost a thousand years. By 1900 B.C. the Babylonians had conquered most of the peoples in the Tigris-Euphrates Valley. About 910 B.C. Babylonia was conquered by the Assyrians.

The Babylonians were the first to make use of the arch in brick buildings. They irrigated their lands and improved the ways of making ditches and canals. They studied the stars and made a calender. Their writing system was very unusual. It was called cuneiform (kū nē'ə form), which means wedge-shaped.

**narrative without time line

THE ASSYRIANS

The Assyrians lived north of Babylonia, on the Tigris River. They developed a civilization by about 2100 B.C. At one time they were a peaceful people and stayed in their own land. Later they became fierce and warlike. With huge armies they swept down on neighboring cities. Armed with iron spears and swords, they struck terror into the hearts of other peoples.

In 910 B.C. the Assyrians conquered Babylonia. By 625 B.C. they had built a huge empire of captured peoples. This empire included not only the people in the Tigris-Euphrates Valley, but the Egyptians as well.

But Assyria was destroyed by invaders from the south and the east.

While the Assyrians were in power they developed trade to distant lands. They built great roads and cities. One of their kings brought together a famous library of clay tablets. They tell us much about the early civilizations.

THE GREEKS

The Greeks lived in a smaller area of land than the other peoples we have discussed. Furthermore, this land was mountainous, rocky, and certainly not fertile. For this reason, many Greeks moved to nearby islands and to fertile coastlands in distant places. Many of them became traders.

Most of the ancient greeks lived in city-states. These were cities which were like small, independent countries in themselves. The Greek people were loyal to their own city-state. There were hundreds of these city-states in ancient Greece.

The ancient Greek civilization began about 1200 B.C. It ended as such about 338 B.C. By 338 B.C. Phillip of Macedonia and his famous son, Alexander the Great, had taken over the Greek city-states. These states had been weakened by fighting and jealousy among themselves. One by one, the city-states lost their independence and became part of Alexander's kingdom.

In spite of their small land area and rocky soil, the ancient Greeks developed a civilization which is most important to us. They were the first people to practice democracy (di mok'rə sē). They were the first to have the Olympic (ō lim'pik) Games. Many of our ideas in the fields of government, literature, and art have come from the ancient Greeks.

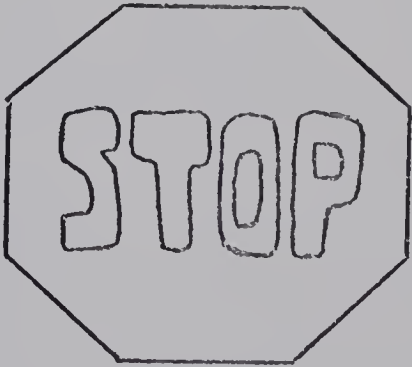
** The phonetic spellings remain unchanged from Livingstone's "Ancient Civilizations" in N.B. Smith, Be A Better Reader, Foundations C, Teacher's Edition (Englewood Cliffs: Prentice-Hall, 1968), pp. 7-9.

B.C. and A.D.

The time covered in the article which you have just read is from 4000 B.C. to the beginning of A.D. Dates extending from B.C. to A.D. are read backward. Thus, 1000 B.C. means 1,000 years before the birth of Christ, and 2000 B.C. means 2,000 years before the birth of Christ.

A.D. dates are read forward. Thus, 500 A.D. means 500 years after the birth of Christ, and 1000 A.D. means 1,000 years after the birth of Christ. The article you read discussed ancient civilization only. So all of your work in the following questions will be with B.C. time.

Please do not write anything inside this box.



APPENDIX B
CHRONOLOGICAL RELATIONSHIPS TEST

By going back to what you have just read answer the following questions under Part 2 of your answer sheet.

1. Which of these ancient nations was civilized at the earliest date we know of?
 - a) Assyria
 - b) Babylonia
 - c) Greece
 - d) Egypt
2. What was that date?
 - a) 338 B.C.
 - b) 606 B.C.
 - c) 3000 B.C.
 - d) 1100 B.C.
 - e) 4000 B.C.
3. Which civilization appeared at the latest date in ancient history?
 - a) Assyria
 - b) Babylonia
 - c) Greece
 - d) Egypt
4. What was that date?
 - a) 3000 B.C.
 - b) 1200 B.C.
 - c) 338 B.C.
 - d) 671 B.C.
 - e) 3500 B.C.
5. Which civilization ended first?
 - a) Assyria
 - b) Babylonia
 - c) Greece
 - d) Egypt
6. What was that date?
 - a) 671 B.C.
 - b) 910 B.C.
 - c) 1900 B.C.
 - d) 338 B.C.
 - e) 606 B.C.
7. Which civilization ended last?
 - a) Assyria
 - b) Babylonia
 - c) Greece
 - d) Egypt
8. What was that date?
 - a) 671 B.C.
 - b) 910 B.C.
 - c) 612 B.C.
 - d) 1200 B.C.
 - e) 338 B.C.
9. Which of the ancient civilizations had the longest life span?
 - a) Assyria
 - b) Babylonia
 - c) Greece
 - d) Egypt
10. About how many years did its life span cover?
 - a) 3,000
 - b) 2,3000
 - c) 2,000
 - d) 1,500
 - e) 1,000
11. Which of the ancient civilizations had the shortest life span?
 - a) Assyria
 - b) Babylonia
 - c) Greece
 - d) Egypt
12. About how many years did its life span cover?
 - a) 1,200
 - b) 2,300
 - c) 300
 - d) 900
 - e) 500

13. Which two civilizations lasted the longest time?
- a) Babylonia, Assyria
 - b) Greece, Egypt
 - c) Assyria, Greece
 - d) Egypt, Babylonia
14. Which two civilizations lasted the shortest time?
- a) Assyria, Greece
 - b) Greece, Babylonia
 - c) Egypt, Assyria
 - d) Babylonia, Egypt

APPENDIX C
INTELLIGENCE QUOTIENT CATEGORIES

EXCERPT FROM THE GUIDE TO INTERPRETATION OF THE CALIFORNIA
TEST OF MENTAL MATURITY SERIES; 1963 REVISION

DESCRIPTIVE CATEGORIES

The table below presents descriptive categories that can be used to evaluate intelligence quotients, percentile ranks, standard scores, and stanines for individual students on factor, section, and total scores of the CTMM series. These relationships are based upon the assumption that the normative population was normally distributed.

A TABLE FOR CLASSIFYING INTELLIGENCE
QUOTIENTS, PERCENTILE RANKS, STANDARD
SCORES, AND STANINES BY A SYSTEM OF
DESCRIPTIVE CATEGORIES

CATEGORY	INTELLIGENCE QUOTIENT RANGE	PERCENTILE RANK RANGE	STANDARD SCORE RANGE	STANINE
Very High	130+	97-99	69+	9
High	116-129	84-96	60-68	7-9
High Average	108-115	70-83	55-59	6-7
Average	93-107	31-69	46-54	4-6
Low Average	85-92	17-30	41-45	3-4
Low	71-84	4-16	32-40	1-3
Very Low	70-	1-3	31-	1

APPENDIX D
DIFFICULTY AND DISCRIMINATION INDICES
FOR THE CHRONOLOGICAL
RELATIONSHIPS TEST

DIFFICULTY OF ITEMS FROM THE CHRONOLOGICAL RELATIONSHIPS TEST*

122

Level of Difficulty	Item
0.670	7
0.651	3
0.632	1
0.599	2
0.585	8
0.517	9
0.495	13
0.493	14
0.425	5
0.417	10
0.417	11
0.415	4
0.403	12
0.354	6

* ordered from least to most difficult.

DISCRIMINATION INDICES OF ITEMS FROM THE CHRONOLOGICAL RELATIONSHIPS TEST**

Discrimination Index (Biserial Correlation)	Item
0.552	2
0.554	12
0.645	7
0.653	4
0.658	8
0.688	3
0.722	5
0.730	14
0.739	6
0.741	10
0.744	11
0.781	1
0.828	13
0.886	9

** ordered from least to most discriminating items.

APPENDIX E

QUESTIONNAIRE AND SUMMARY OF
QUESTIONNAIRE COMMENTS

QUESTIONNAIRE ON TEACHING OF CHRONOLOGY

Teacher _____ School _____

1. Which of these ancient civilizations have your students studied this year? (Check those you have covered)

<input type="checkbox"/> Babylonia	<input type="checkbox"/> Sumeria	<input type="checkbox"/> Rome
<input type="checkbox"/> Mesopotamia	<input type="checkbox"/> Egypt	
<input type="checkbox"/> Assyria	<input type="checkbox"/> Greece	

If there are any other ancient civilizations which your class has studied this year list them here.

2. Have you taught anything about time or chronology this year? _____ YES _____ NO

If yes, please write a brief summary of what you did.

3. Do you consider time lines in reference or text books to be useful? _____ YES _____ NO
If yes, how are they useful? If no, why do you think they are not useful?

4. Have you given any instruction in the construction and use of time lines this year? _____ YES _____ NO
If yes, please write a brief summary of what you did.

Of the five teachers who were teaching Grade 6 Social Studies all replied to the questionnaire. Their answers and comments are reported in the summary below.

Three teachers taught units on Ancient Egypt and two taught on Ancient Rome. No teacher taught a unit which included any of the other ancient civilizations.

Only one of the five had taught something about time or chronology. That person constructed a time line contrasting the several hundred years of Rome's "glory" with the short period of time covered by the Space Age.

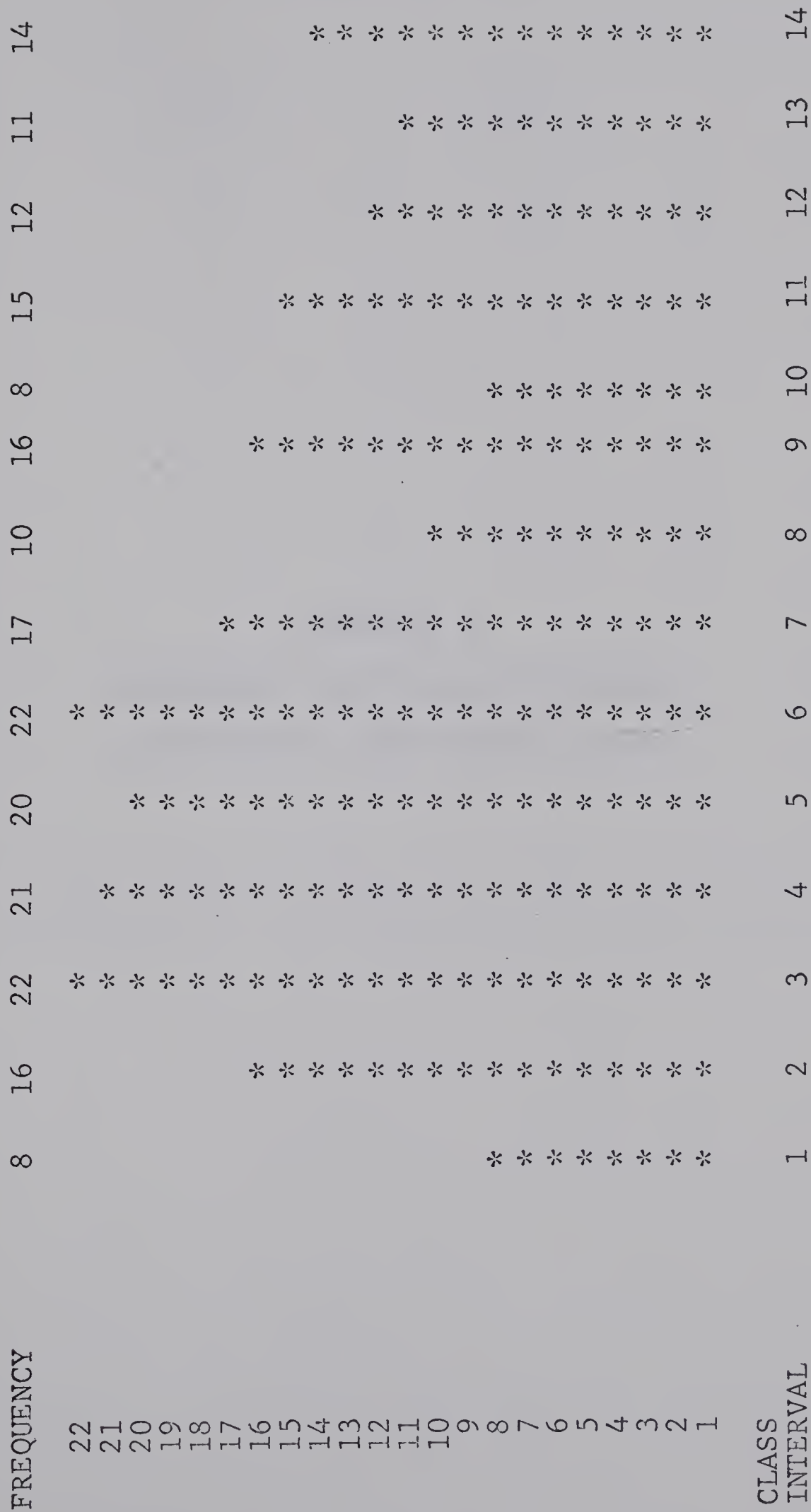
Two of the five teachers thought that time lines in books were not useful because students in grades 5 and 6 cannot easily read time lines. However, the three other teachers thought that such time lines were useful by providing a comparison for individual dates, or showing an isolated incident in relation to a whole.

Only one of the five teachers had given instruction in the construction and use of time lines. She had constructed a time line covering the period from the beginning of Rome's importance to the decline and fall of the Western Roman Empire.

APPENDIX F

HISTOGRAM OF CHRONOLOGICAL
RELATIONSHIPS TEST SCORES

HISTOGRAM SHOWING DISTRIBUTION OF CHRONOLOGICAL RELATIONSHIPS TEST SCORES



Mean = 7.04
Median = 6.5

KR-20 Reliability = 0.78
Variance = 14.68

APPENDIX G

HOMOGENEITY OF VARIANCE BETWEEN
EXPERIMENTAL AND CONTROL GROUPS

HOMOGENEITY OF VARIANCE BETWEEN EXPERIMENTAL AND CONTROL GROUPS

INDEPENDENT VARIABLE LEVEL	EXPERIMENTAL GROUP VARIANCE	N	CONTROL GROUP VARIANCE	N	P*
High I.Q.	15.65	36	14.06	28	.78
High Ave. I.Q.	13.87	27	12.35	34	.74
Ave. I.Q.	11.91	25	5.84	26	.08
Low I.Q.	5.61	10	6.14	14	.91
High Reading	14.02	49	11.27	46	.47
Ave. Reading	12.73	30	10.04	38	.57
Low Reading	8.72	21	4.08	22	.09
Boys	16.78	69	11.66	56	.16
Girls	13.09	37	10.62	51	.49

* probability resulting from performing F-tests on the pairs of variances.¹ The P value indicates the probability that the two variances are homogeneous (from samples drawn from the same population). None of the values are significant at the .05 level.

¹ see discussion in Weinberg and Schumaker, op.cit., pp. 209-304.

APPENDIX H
DIFFERENCES BETWEEN EXPERIMENTAL
AND CONTROL GROUPS
BY I.Q. LEVEL

TESTS FOR SIGNIFICANT DIFFERENCES BETWEEN EXPERIMENTAL AND CONTROL GROUP
 MEAN CHRONOLOGICAL RELATIONSHIPS TEST SCORES AT EACH I.Q. LEVEL

I.Q. Level	Experimental Group		Control Group		Difference Between Means		t	DF	Probability	Significance*
	N	Mean	N	Mean	Mean	Mean				
High (>115)	36	9.94	28	7.29	2.65	2.667	61	.004**		S
High Ave. (108-115)	27	8.59	34	6.35	2.24	2.368	59	.02		S
Ave. (93-107)	25	6.08	26	4.92	1.16	1.363	49	.18		NS
Low (<93)	10	4.70	14	4.66	.04	0.051	22	.96		NS***

* S - probability of .05 or less.
 NS - probability of greater than .05.
 ** - using a one-tailed test for reasons explained in chapter 3.
 *** - The t-test was used here for consistencey. However, since the sample included less than 30 subjects, the Median Test, which is the non-parametric test corresponding to the parametric T-test, was used to check the result. The Median Test yielded a probability of greater than .30, thus supporting the result of the t-test.

APPENDIX I
DIFFERENCES BETWEEN EXPERIMENTAL
AND CONTROL GROUPS BY
READING COMPREHENSION
LEVEL

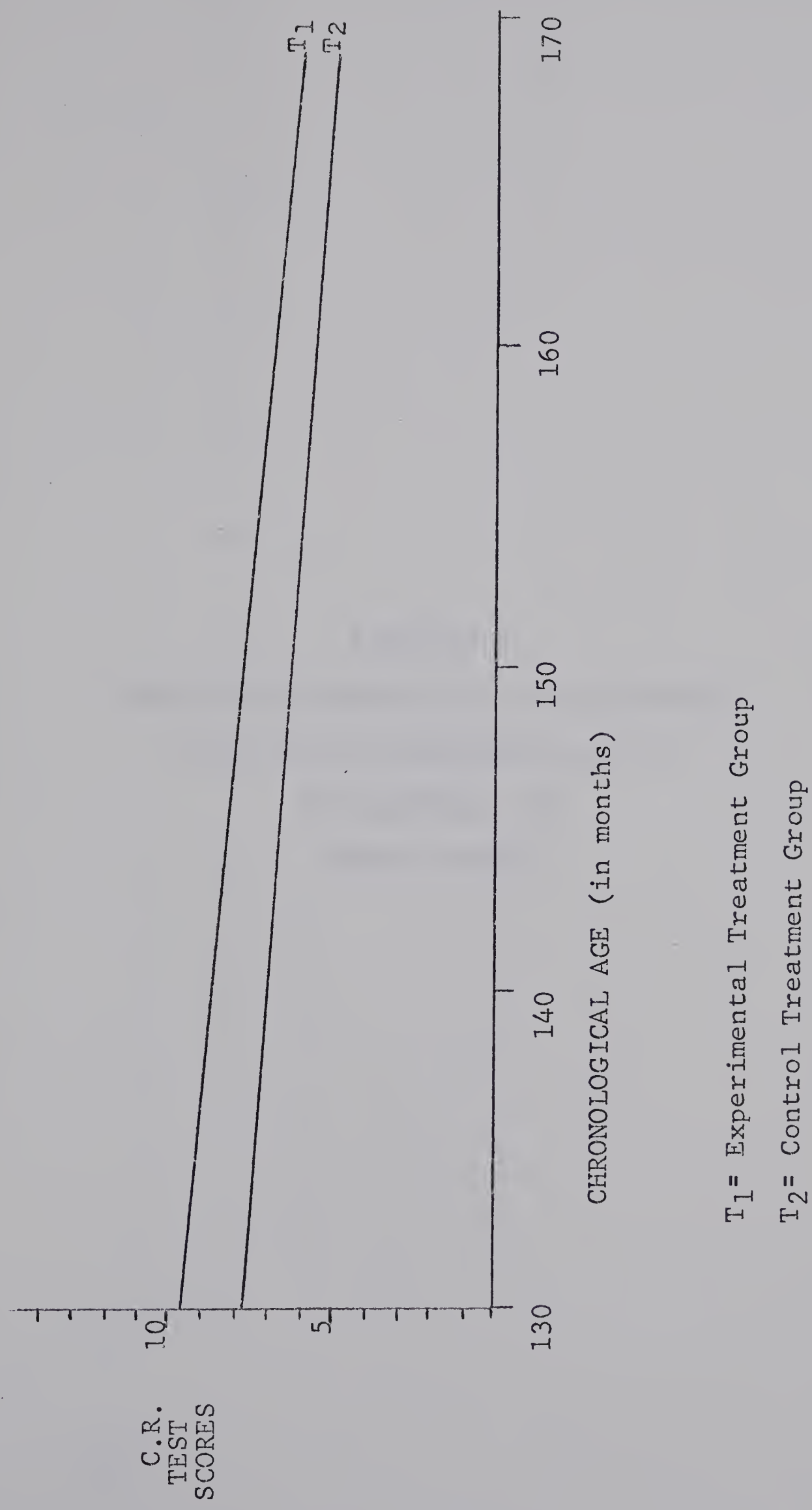
TESTS FOR SIGNIFICANT DIFFERENCES BETWEEN EXPERIMENTAL AND CONTROL GROUP MEAN
 CHRONOLOGICAL RELATIONSHIPS TEST SCORES AT EACH READING
 COMPREHENSION LEVEL

Reading Comprehension Level	Experimental Group		Control Group		Difference Between Means		t	DF	Probability	Significance*
	N	Mean	N	Mean	Mean	Means				
High (>Gr. 6.9)	49	9.98	46	7.53	2.45	3.289	92	.001		S
Ave. (Gr. 6.0-6.9)	30	7.34	38	5.47	1.87	2.255	65	.03		S
Low (<Gr. 6.0)	21	5.43	22	3.77	1.66	2.103	41	.02		S

* S - probability of .05 or less.
 NS - probability of greater than .05.

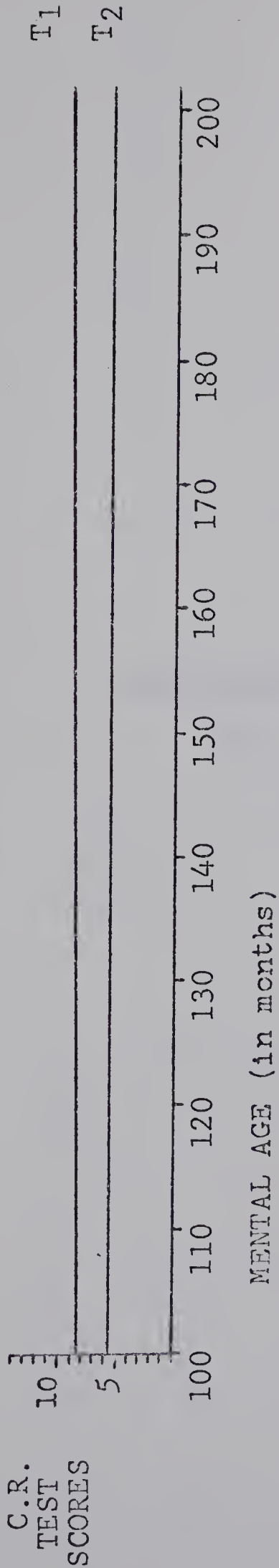
APPENDIX J
PREDICTING CHRONOLOGICAL RELATIONSHIPS
TEST SCORES FROM CHRONOLOGICAL AGE
IN EXPERIMENTAL AND
CONTROL GROUPS

PREDICTING CHRONOLOGICAL RELATIONSHIPS TEST SCORES FROM CHRONOLOGICAL AGE IN EXPERIMENTAL AND CONTROL GROUPS



APPENDIX K
PREDICTING CHRONOLOGICAL RELATIONSHIPS
TEST SCORES FROM MENTAL AGE IN
EXPERIMENTAL AND
CONTROL GROUPS

PREDICTING CHRONOLOGICAL RELATIONSHIPS TEST SCORES FROM MENTAL AGE IN EXPERIMENTAL AND CONTROL GROUPS



T₁= Experimental Treatment Group
T₂= Control Treatment Group

APPENDIX I
DIFFERENCES BETWEEN EXPERIMENTAL
AND CONTROL GROUPS FOR
BOYS AND GIRLS

TESTS FOR SIGNIFICANT DIFFERENCES BETWEEN EXPERIMENTAL AND CONTROL GROUP
MEAN CHRONOLOGICAL RELATIONSHIPS TEST SCORES FOR
BOYS AND GIRLS

Sex	Experimental Group		Control Group		Difference Between Means	t	DF	Probability	Significance*
	N	Mean	N	Mean					
Boys	69	8.66	56	6.36	2.30	3.330	122	.001	S
Girls	37	7.22	51	5.65	1.57	2.104	86	.04	S

* S - probability of .05 or less.
NS - probability of greater than .05.

B29963